Sequence Specific Recognition and Cleavage of DNA by Synthetic Molecules

Thesis by

R. Scott Youngquist

In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

1988

(submitted August 5, 1987)

to God and my family

Acknowledgments

I would like to thank my advisor, Peter Dervan, for providing the research opportunities which made this thesis possible and for his enthusiastic guidance and support throughout my career at Caltech.

I am also indebted to Jim Sluka for critical contributions and expert advice. I am grateful to Peter Schultz, Burt Leland, Keith Harshman, John Griffin and the other members of the Dervan group, past and present, for scientific discussions as well as critical distractions. I would also like to thank John Griffin, Alanna Shepartz, and Warren Wade for proofreading this thesis and contributing constructive comments.

Abstract

DNA recognition is an essential biological process responsible for the regulation of cellular functions including protein synthesis and cell division and is implicated in the mechanism of action of some anticancer drugs. Studies directed towards defining the elements responsible for sequence specific DNA recognition through the study of the interactions of synthetic organic ligands with DNA are described.

DNA recognition by poly-*N*-methylpyrrolecarboxamides was studied by the synthesis and characterization of a series of molecules where the number of contiguous *N*-methylpyrrolecarboxamide units was increased from 2 to 9. The effect of this incremental change in structure on DNA recognition has been investigated at base pair resolution using affinity cleaving and MPE·Fe(II) footprinting techniques. These studies led to a quantitative relationship between the number of amides in the molecule and the DNA binding site size. This relationship is called the $n + 1$ rule and it states that a poly-*N*-methylpyrrolecarboxamide molecule with n amides will bind $n + 1$ base pairs of DNA. This rule is consistent with a model where the carboxamides of these compounds form three center bridging hydrogen bonds between adjacent base pairs on opposite strands of the helix. The poly-*N*-methylpyrrolecarboxamide recognition element was found to preferentially bind poly dA·poly dT stretches; however, both binding site selection and orientation were found to be affected by flanking sequences. Cleavage of large DNA is also described.

One approach towards the design of molecules that bind large sequences of double helical DNA sequence specifically is to couple DNA binding subunits of similar or diverse base pair specificity. Bis-EDTA-distamycin-fumaramide (BEDF) is an octaamide dimer of two tri-*N*-methylpyrrolecarboxamide subunits linked by fumaramide. DNA recognition by BEDF was compared to P7E, an octaamide molecule containing seven consecutive pyrroles. These two compounds were found to recognize the same sites on pBR322 with approximately the same affinities demonstrating that fumaramide is an effective linking element for *N*-methylpyrrolecarboxamide recognition subunits. Further studies involved the synthesis and characterization of a trimer of tetra-*N*-methylpyrrolecarboxamide subunits linked by β-alanine ($(P4)_3E$). This trimerization produced a molecule which is capable of recognizing 16 base pairs of A·T DNA, more than a turn and a half of the DNA helix.

DNA footprinting is a powerful direct method for determining the binding sites of proteins and small molecules on heterogeneous DNA. It was found that attachment of EDTA·Fe(II) to spermine creates a molecule, SE·Fe(II), which binds and cleaves DNA sequence neutrally. This lack of specificity provides evidence that at the nucleotide level polyamines recognize heterogeneous DNA independent of sequence and allows SE·Fe(II) to be used as a footprinting reagent. SE·Fe(II) was compared with two other small molecule footprinting reagents, EDTA·Fe(II) and MPE·Fe(II).

Table of Contents

Chapter 1

Introduction

DNA recognition is an essential biological process responsible for the regulation of cellular functions including protein synthesis and cell division.[1,2] The mechanism of action of some anticancer drugs is also believed to involve sequence specific DNA complexation.[3] Although this is an important area of interest in molecular biology, the particular combination of hydrogen bonding, van der Waals contacts and electrostatic interactions required for sequence specific recognition of DNA are not well understood. The work that I describe here are studies directed towards defining the important elements responsible for specific DNA recognition through the study of the interactions of synthetic organic ligands with DNA.

Distamycin A and Netropsin. Distamycin A[4] and netropsin[5] are oligo-*N*-methylpyrrolecarboxamide antibiotics isolated from *Streptomyces distallicus* and *Streptomyces netropsis* respectively. These compounds exhibit antifungal, antimitotic and antiviral properties *in vivo*.[6-8] They are believed to exert their biological effects *via* the formation of a tightly bound complex with double stranded DNA, possibly in promoter regions.[9]

The molecular structure of each of these compounds has been determined by chemical analysis[10,11] and verified by total synthesis.[12,13] Both drugs are optically inactive; however, x-ray analysis of the sulfate salt of netropsin[14] and a precursor to distamycin[15] show that there is a helical twist of 4° or 26° between adjacent pyrrole rings creating asymmetric conformations for the two molecules. Both netropsin and distamycin are crescent-shaped with the amide hydrogens

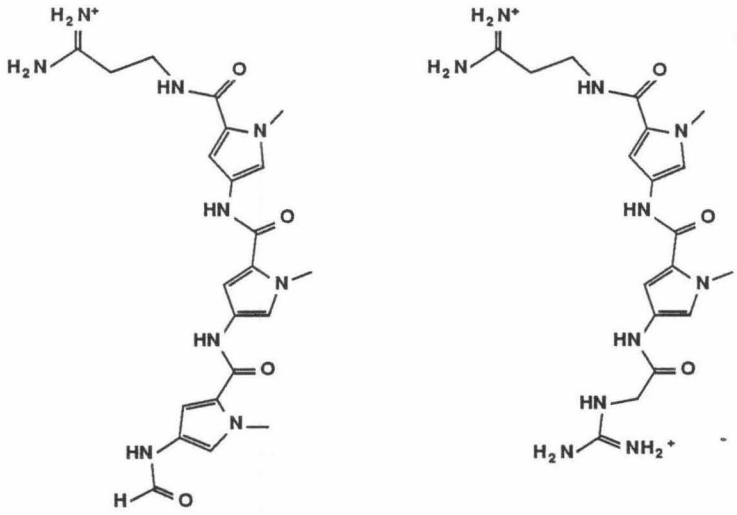


Figure 1.1 Distamycin A and Netropsin

on the concave side and the carbonyls and *N*-methyls on the convex face (fig. 1.1).

Extensive biophysical studies on the interaction of distamycin A and netropsin with DNA have been carried out. These studies indicate that the molecules bind to A·T rich regions of DNA. This work has been extensively reviewed in a recent report by Zimmer.[16] The discussion here will concentrate on the nature of the *N*-methylpyrrolecarboxamide-DNA interaction and its effect on flanking sequences.

Effects on DNA Structure. In addition to a high preference for A·T sequences, netropsin and distamycin have been found to be specific for B-form DNA. DNA in ethanolic solution may be converted to A-form by a lowering of

the water activity.[17] Either distamycin or netropsin reverts calf thymus DNA to B-form even in 82% ethanol.[17, 18] Polymers containing poly 2′-deoxy-2′-flouro-adenylic acid are known to prefer the A-form in aqueous solution, but they too are converted to B-form by either drug.[19, 20] The same effect is seen with poly(dI)·poly(2′-deoxy-2′-flouro dC).[20] Additional evidence for the B-form preference is the lack of binding to double stranded RNA or DNA-RNA hybrids both of which are known to exist in the A-form.[8, 19, 21, 22]

In addition to the ability to change the helical form of DNA, netropsin and distamycin can affect the structure of DNA. DNA viscosity measurements with distamycin and netropsin reveal differences in drug binding. Netropsin produces an increase in both the contour length and the persistence length of DNA.[23, 24] The persistence length increase is caused by stiffening of the DNA helix which freezes out some of the degrees of freedom. This effect was found to be dependent on cationic counterions.[25] The contour length is increased by changing the DNA to a more extended configuration and was found to be counterion independent. Distamycin causes no change in contour length, but displays a persistence length characteristic of bending or kinking.[26–28] Calculations by Reinert concluded that the distamycin recognition pattern is about 0.01 nm larger than the DNA pattern.[28] The original viscosometric titration and sedimentation coefficient studies showed no change in DNA winding by either netropsin or distamycin.[21, 29] Using a more sensitive technique, Malcolm and Snovnou have recently shown that netropsin winds DNA from plasmid pAT 153 approximately

+7.5° per bound netropsin while distamycin has either a small or no effect on the same plasmid.[30]

The use of magnetic circular dicroism (MCD) permits the observation of the effect of the drugs on the circular dicroism (CD) of DNA.[31,32] From these studies, radical changes in the tilt of the bases with respect to the helical axis upon distamycin or netropsin binding have been ruled out. Crothers *et al.* used electric dicroism studies to confirm these results and in addition show that the DNA is not substantially kinked or bent under 2.5 mM salt conditions.[33] Under the same low salt conditions as the electric dicroism studies, the Crothers group saw a very large increase in apparent DNA length of 14 Å per drug bound with either netropsin or distamycin using two eukaryotic DNAs, calf thymus and chicken erythrocyte.[33,34] However, under the same conditions another eukaryotic DNA, human placenta, did not show the large length increase nor did prokaryotic *E. Coli* DNA or the homopolymer poly(dA)·poly(dT). Because all the natural DNAs had approximately the same A·T content, the phenomenon was considered source dependent and independent of base content.

Kinetoplast DNA is known to exist in a kinked form which arises from a repetitive sequence of A·T rich DNA. Crothers *et al.* have shown that distamycin will bind to this fragment and remove those kinks causing the fragment to run more slowly in an acrylamide gel.[35] In the gel demonstrating this effect, distamycin also causes a fragment from pBR322 to run faster. This effect is consistent with the poly-*N*-methylpyrrolecarboxamide compound binding the

DNA and *inducing* bending. This is a clear example of distamycin producing changes in the structure of the DNA and, because it occurs on only a few of the restriction fragments, establishes that this effect is sequence dependent.

DNA Complex Structure Studies. Recent crystal structure data on the complex of netropsin with a B-form dodecamer, 5′–CGCGAATTCGCG–3′, provides a "snapshot" of the recognition of DNA by *N*-methylpyrrolecarboxamides.[36–38] Netropsin sits symmetrically in the center of the minor groove of the right-handed DNA displacing the water molecules which had formed the spine of hydration. Each of the amide groups forms a bridged hydrogen bond with the adjacent adenine N-3 or thymine O-2 atoms on opposite strands of the helix. Netropsin binding does not induce any major structural changes in the DNA neither unwinding nor elongating the dodecamer. It does, however, force open the minor groove by 0.5 to 2.0 Å and bend the helix axis back by 8° across the region of attachment. Dickerson and co-workers suggest that the sequence specificity of netropsin arises not from the bridging hydrogen bonds, but from close van der Waals contacts between adenine C-2 hydrogens and the CH-3 group on the pyrrole ring. Additionally, the guanine 2-amino group protruding from the floor of the minor groove obstructs efficient recognition of G·C regions by netropsin.

The netropsin-DNA co-crystal has provided valuable insight into the mechanism of DNA recognition by *N*-methylpyrrolecarboxamides; however, there are some important points to consider regarding a quantitative extrapolation of

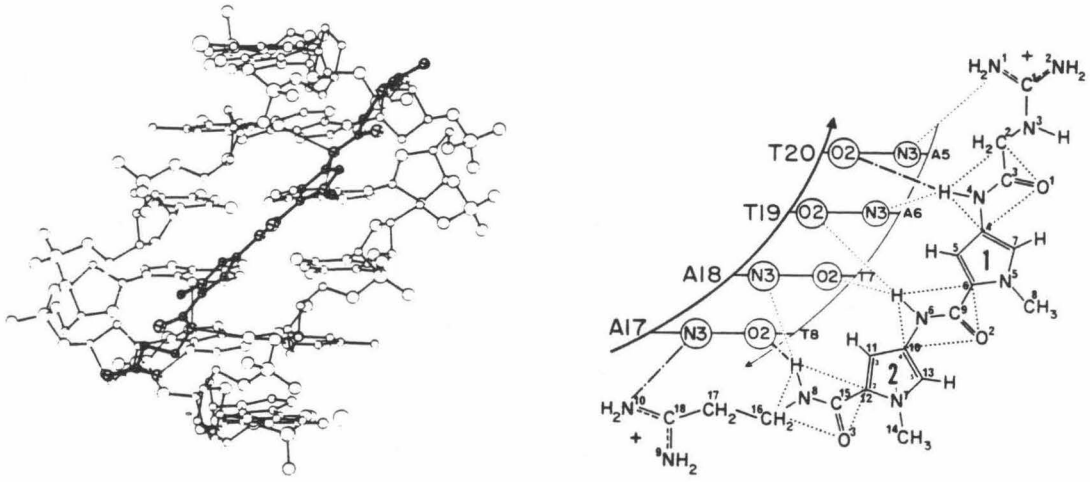


Figure 1.2 X-tal Graphics

(Left) Ball and stick model of the netropsin-DNA crystal structure data. Open circles are DNA atoms; crossed circles are netropsin atoms. (Right) Diagramatic representation of netropsin binding to the minor groove of the DNA. Dot-dashed lines represent distances short enough to be standard hydrogen bonds. Dotted lines indicate distances of 3.2Å or more (from ref. 37).

these data from the crystal to the solution phase. In the crystal, the terminal three G·C base pairs of each duplex form five H-bonds with atoms of the next duplex above and below in a manner that requires a bend in the helix.[36–38] This interaction effectively locks the base pair helical repeat, eliminating any effect that netropsin might have on this structural variable. There is also evidence from model building based on NMR and Raman data that the minor groove in this crystal structure may be anomalously narrow implying that netropsin does not *necessarily* have to sit symmetrically in the center of the groove.[39, 40] Hydrogen bond distances from the crystal structure itself show a much tighter

interaction between the amides of the netropsin and the thymine O2 than with the adenine N3 indicating that the molecule may prefer to associate with the thymine residues (table 1.1). The structural features of the crystal structure may also have been influenced by the high concentration (33%) of 2-methyl-2,4-pentanediol used to induce crystal formation.

Table 1.1: Crystal Structure H-bond Distances[37]

Netropsin atom	DNA atom	Distance Å
N-4	O-2 of T-20	2.68
N-4	N-3 of A-18	3.28
N-6	O-2 of T-19	3.86
N-6	O-2 of T-7	3.54
N-8	N-3 of A-18	3.45
N-8	O-2 of T-8	2.56

Atom numbers refer those in figure 1.3 (ref. 37)

Studies directed at determining the optimal sequence for DNA recognition by *N*-methylpyrrolecarboxamides suggest that netropsin preferentially recognizes poly(dA)·poly(dT) sequences of at least four base pairs; however, many sites appear to be longer sequences of five or more bases contained in longer stretches of A·T DNA.[41–46] These results argue that the sequence being recognized by netropsin in this crystal structure may not be the one most efficiently recognized in solution and; therefore, the quantitative results may not be applicable to solution complexes. A final point is that predictions made from a quantitative interpretation of the crystal structure have been inaccurate: the estimate of the largest poly-*N*-methylpyrrolecarboxamide chain permissible for

DNA recognition was incorrect (*vide infra*) as was the prediction that replacement of pyrrole by imidazole would result in a G·C specific molecule.[47]

Other structural data on the *N*-methylpyrrolecarboxamide-DNA complex has been aquired from NMR studies. The earliest NMR work with netropsin:DNA complexes was carried out by Patel *et al.* using homopolymer poly d(AT) DNA.[48] From chemical shift differences, they determined that at a phosphate to drug (P/D) ratio of 50, netropsin not only stabilized the region of DNA that it bound, but also lowered the rate of melting in antibiotic free regions by 10-fold with respect to native poly d(AT). Of the DNA protons, the deoxyribose A(H-1′)s and T(H-1′)s showed the the largest chemical shift difference upon formation of the netropsin-DNA complex. These shifts represent netropsin induced changes in drug free regions of the DNA consistent with changes in the glycosidic tortional angles. The observation that netropsin binding changes both the rate of melting and the glycosidic torsional angles in drug free regions of DNA demonstrates that netropsin affects DNA structure remote from its binding site.

Patel and co-workers then began to study oligonucleotides of defined sequence to investigate the netropsin-DNA interaction in more detail. Using the oligonucleotide $d(CGCGAATTCGCG)_2$, Patel *et al.* employed imino exchange experiments to confirm that netropsin stabilizes the DNA duplex including the G·C residues adjacent to the netropsin binding site.[49] They also found that this effect is not due to a change in the activation energy for base pair opening.

With a series of NOE experiments, Patel and co-workers found distances between the netropsin H-3 and adenine H-2 of 2.5 Å.[50] These data are consistent with netropsin binding in the minor groove with the amides directed towards the DNA helix.

Recent work by Patel *et al.* has focused on the interaction of netropsin with the oligonucleotides $d(GGTATACC)_2$ and $d(GGAATTCC)_2$.[51-54] In these studies, Patel and co-workers found that the DNA remains right-handed and that the DNA conformation with netropsin bound is similar to that for the DNA free in solution. Nine contacts between netropsin and the minor groove were detected using two dimensional NOE experiments verifying that netropsin binds in the minor groove with approximately the same geometry seen in the crystal structure. Some of the cross peaks expected for B-form DNA were missing in the 2D NOE experiments on the netropsin-DNA complexes indicating that small conformational changes occur at the binding site possibly extending into the G·C flanking regions. The ^{31}P NMR spectrum showed changes in the phosphorus resonances indicative of a change in the DNA backbone conformation with netropsin binding.

The structural data for the two oligomer-DNA complexes are similar; however, there are some specific differences. The imino exchange rate for the d(TATA) and d(AATT) cores are not the same. The d(TATA) oligo shows rapid proton exchange determined by the loss of peak doubling at 35°C while the d(AATT) oligo does not show this effect until ~50°C. The chemical shift

of the netropsin amide protons upon complexation with the two DNAs are also different. With the d(TATA) core, the amide peaks are shifted to lower field than with the d(AATT) core. This is consistent with stronger H-bonding with d(TATA); however, the exchange rate data and results by other workers on the stability of netropsin:DNA complexes has shown that the alternating A·T site is a poorer netropsin binding site. Patel and co-workers interpret this result as demonstrating that factors other than hydrogen bonding are important for stabilizing the netropsin-DNA complex.

Reid *et al.* have used NMR to study the interaction between distamycin A and a dodecamer of DNA $d(CGCGAATTCGCG)_2$.[55] Binding the asymmetric distamycin molecule to the symmetric DNA destroys the DNA's C_2 symmetry which results in a doubling of peaks in the NMR spectrum. Negative NOEs for the distamycin molecule show that it is tumbling with about the same correlation time as the much larger DNA molecule which means that the distamycin is very tightly bound to the DNA. The absence of distamycin·distamycin cross-peaks in the two-dimension exchange experiments is consistent with nearly all the drug being bound to DNA with a single binding mode. As the temperature is increased from 27°C, distamycin begins to exchange between DNA molecules causing the spectrum to broaden.

The strongest NOEs observed with the non-exchangable protons in the complex are between the pyrrole H3's and adenine C2H's, identifying these as the points of nearest contact between drug and DNA. There is one-to-one correspondence between the pyrrole rings 1, 2, and 3 and the consecutive adenines

A5, A6, and A18, respectively, in the complex. This correlation is not seen with A6, A18, and A17, another 3 consecutive adenines, indicating that this alternate binding site must consitute less than 10% of the total complex. The NOE experiments confirm that distamycin binds in the minor groove with the amides directed towards the DNA and that of two possible binding sites distamycin predominantly recognizes only one.

Some of the expected NOEs for B-DNA were missing in these studies consistent with drug binding altering the conformation of the DNA slightly; however, only the resonances from A5 through T7 give the double exchange peaks indicative of major structural alteration. The resonances from the residues near the ends of the dodecamer are nearly superimposable before and after distamycin binding showing that they are little perturbed from the original structure. Some NOE measured distances in the complex are slightly changed from those measured in the netropsin-DNA co-crystal. The NOE measured distances between T20 methyl and C21 CH-5 and CH-6 are 4 Å instead of the 5.0 Å and 6.8 Å, respectively, measured from the crystal structure. The consequence of such a change is a greater separation between the edges of the two bases in the minor groove than in the solution phase structure.

Both free and bound forms of distamycin are found in the NMR spectrum; therefore, exchange must be slow on the NMR timescale. The limiting behavior for slow-exchange from a chemical shift standpoint is $k_{ex}=|\Delta\omega|$. At 27°C, the smallest chemical shift difference between the free and bound forms is 15 Hz

(for the methyl resonance of T8). Since the binding constant of distamycin, K_{app}, is around 10^7 M^{-1}, k_{on} is much greater than k_{off}. This means that k_{ex} is essentially a measure of k_{off}, and must be slower than 15 s^{-1}. From diagonal peak intensities in the NOE, k_{off} can be calculated to be $\sim$4 s^{-1}; however, this number does not correct for relaxation processes that may have occurred during the mixing time. Another set of cross-peaks for the DNA is consistent with drug binding but no cross-peaks are observed for the distamycin. These peaks can be explained by "flip-flopping" of the distamycin on the binding site. The peak intensities were analyzed to give a measure of k_{flip} as $\sim$2 s^{-1} at 27°C.

Thermodynamics. Thermodynamic data on the netropsin/distamycin interaction with DNA have been collected through binding constant and calorimetric studies. The first determination of the binding constant of netropsin to DNA was carried out by Wartell *et al.* using two standard techniques, equilibrium dialysis and spectral titration. The values they reported were between 10^5 M^{-1} and 10^6 M^{-1} with a variety of DNAs with a buffer system of 0.1 M NaCl, 1×10^{-3} M phosphate, pH 6.4.[56] These standard methods of binding constant determination are complicated by both the instability of the compound in solution and the small amount of free drug at low concentration. Equilibrium dialysis assumes no netropsin degradation during the process and suffers from the small amount of unbound drug at low or even moderate DNA saturation. Spectral titrations are difficult because the $\Delta\epsilon$ of bound *vs.* unbound drug is small and this method must assume a linear relationship between optical change and extent of binding.

To circumvent these difficulties, McGhee used a method of fitting theoretical DNA melting curves to experimental ones with netropsin and poly d(AT).[57] This approach is rapid, limiting netropsin degradation, and does not depend on the amount of unbound drug. From these data, he determined a binding site size of 4 base pairs and a binding constant of 5×10^8 M^{-1} for netropsin. The data were analyzed using an equation derived by Crothers.[58]

$$\frac{1}{T_m^o} - \frac{1}{T_m} = \frac{R}{\Delta H} \ln \left[\frac{(1 + K_h L)^{1/n_h}}{(1 + K_c L)^{2/n_c}} \right] \qquad (1.1)$$

T_m^o and T_m are the Kelvin temperatures corresponding to the midpoints of the helix-to-coil transition in the absence and presence of netropsin, respectively. ΔH is the enthalpy change for the melting of a base pair in the absence of bound ligand. K_c and K_h are the binding constants to coil and helix forms respectively. n_c and n_h are the ligand size for coil and helix binding respectively. L is the free ligand activity. The curve fitting assumes that the drug does not bind single-stranded DNA and that binding is not cooperative. Thus, the only two adjustable parameters in eq. 1.1 are n_h (size in base pairs) and K_h (binding constant). These values were determined in two ways. First n_h was determined by the slope of the melting curve which most closely matched that of the DNA-only melting curve. According to theory, this match should occur at 200% saturation and in this study gave a value of 3.8 bp per bound netropsin. With n_h and equation 1.1, the binding constant was determined to be 4×10^8 M^{-1}. The second method McGhee used to determine n_h was the generation of theoretical curves for the half-saturated case and comparison to experimental data. The

values producing the best fit were $n_h = 4$ and $K_h = 5\times10^8$ M^{-1}. McGhee noted from his calculations that the height of the first melting phase depended almost solely on the binding site size and was nearly independent of the binding constant. The buffer concentrations for these binding studies measurement was 4×10^{-3} M cacodylic acid, 2×10^{-3} M NaOH, 1×10^{-4} M Na_2EDTA, pH 6.25.

The co-sedimentation of distamycin or netropsin with calf thymus DNA was used by Luck *et al.* to determine their respective binding constants.[59] This method assumes that all binding sites are independent and equivalent. Their data did not go below r = 0.05 (r = number moles of antibiotic bound per mole (total) DNA phosphate) to allow determination of a binding constant for the highest affinity sites because of the limited sensitivity of their scanning system. The data were analyzed using the equation:

$$\frac{[A]}{r} = \frac{\alpha}{K} + \frac{\alpha}{[A]} \tag{1.2}$$

Where [A] equals the molar concentration of nonbound ligand molecules, α is the number of DNA phosphates per binding site and K is the association constant. Best fitting binding curves for the original data were calculated, assuming that the equilibrium is governed by the law of mass action for a single class of independent binding sites. This simplification avoided the problem of having points with relatively high experimental error disproportionately weighted. For netropsin the best fit was obtained with $\alpha = 6.0$ and $K = 2.9\times10^5$ M^{-1}. With distamycin the values were $\alpha = 6.1$ and $K = 11.6\times10^5$ M^{-1}. These measurements were carried out in 0.05 M tris·HCl, pH 7.8. Using data published by

Mazza *et al.*[60] with a ^{14}C-labeled distamycin and SPP1 phage DNA in a preparative sucrose gradient, Luck *et al.* calculated the high affinity binding constant as 2.4×10^{9} M^{-1} with an α of 16.1 nucleotides. The r for this measurement was $\approx$ 0.06 with a free distamycin concentration of about 10^{-8} M. The Scatchard analysis indicated one type of binding site. No information was given about the buffer conditions.

Gursky *et al.* used dansyl derivatives of distamycin to measure binding constants by fluorescence.[61,62] These studies have the advantage of the high sensitivity afforded by the dansyl probe; however, the attachment of this probe to the molecule almost certainly influences the efficiency of polypyrrole binding to DNA. However, because the dansyl group attachment did not appreciably affect the CD spectrum of the modified molecules compared with the natural antibiotic, they assumed that dansyl attachment did not affect the geometry of binding. Scatchard analysis of the titrations gave the following binding constants: 5.5×10^{7} M^{-1} for poly(dA)·poly(dT), 1.8×10^{7} M^{-1} for poly d(AT), 5×10^{4} M^{-1} for poly dG·poly dC, and 2×10^{4} M^{-1} for poly d(AC)·poly d(TG). Assuming that the pyrrole interacts with only one chain of the double helix, they calculated the free-energy change for each base in a binding site (table 1.2).

Gursky *et al.* also carried out experiments with a homologus series of dansyl polypyrrole compounds containing between one and three pyrrole units.[61] In the case of poly(dA)·poly(dT) DNA, the relationship between the free energy of binding and the number of amides in the molecule was a linear one with each

Table 1.2: *N*-methylpyrrole:DNA Stablility constants.[61, 62]

Base	Stability constant	Free-energy change (kcal/M)
Thymine	$S_T = 32 \pm 6$	Δ = -2.1 $\pm$ 0.1
Adenine	$S_A = 5.9 \pm 2$	Δ = -1.1 $\pm$ 0.2
Cytosine	$S_C = 4.2 \pm 1.5$	Δ = -0.9 $\pm$ 0.2
Guanine	$S_G = 0.2^{+1}_{-0.2}$	$\Delta = +0.9\ ^{-1}_{+\infty}$

amide contributing ~2 kcal/mol. This means that individual pyrrole amides contribute an equal amount to DNA binding, consistent with the formation of a new hydrogen bond. While the values determined for poly d(AT)· poly d(AT) were close to those for poly(dA)·poly(dT), they were not the same indicating that the two polymers are recognized similarly, but not identically.

The Crothers group employed a the phase partition method developed by Waring *et al.*[63] to measure the binding constant for distamycin to calf thymus DNA.[33, 34] This method involves partitioning the drug between an aqueous and an organic phase providing the two advantages of rapid equilibration time (~ 2 hours) and amplification of the free drug concentration in the organic phase provided the organic/aqueous partition coefficient is large. The binding constants were determined by calculating the concentration of free drug using equation 1.3. These results were checked by measuring the total distamycin present by dissociating the drug-DNA complex with an equal volume of DMSO and then measuring the distamycin concentration using an extinction coefficient determined under these mixed solvent conditions.

$$c_f = \frac{c_{org}}{K} \tag{1.3}$$

Where K is the partition coefficient and c_f and c_{org} are the concentrations of free drug in the aqueous and the organic phases respectively. The amount of drug bound is calculated using: $c_b = c_a - c_f$ where c_b is the amount of drug bound and c_a is the total amount of distamycin. Crothers and co-workers reported binding affinities for calf thymus DNA and found allosteric binding with high positive cooperativity at low r values and noncooperative binding at r values above 0.03. These results are consistent with the DNA length increases observed with electric dicroism at varying distamycin concentrations. The values determined for the allosteric binding of distamycin to calf thymus DNA are 0.95×10^5 M^{-1} for the initial low affinity sites and 0.95×10^6 M^{-1} for the high affinity sites. The buffer used in these experiments was: 1 M NaCl, 6 mM Na_2HPO_4, 2 mM NaH_2PO_4, 1 mM EDTA, pH 6.5.[33] The distamycin-calf thymus DNA binding constants were also measured under low salt conditions giving 3×10^5 M^{-1} and 4.8×10^6 M^{-1} respectively for the low and high affinity sites (50 mM NaCl, 6 mM Na_2PO_4, 2 mM NaH_2PO_4, 1 mM EDTA, pH 6.5).[34] With *E. Coli* DNA under low salt conditions, no allosteric binding was observed. The main disadvantage of this method is that the measurements are made in a buffer saturated with organic solvent not in aqueous solution. *n*-Pentanol was used as the organic phase.

Marky *et al.*[64,65] used temperature-dependent UV absorbance to determine the binding constant of netropsin to poly(dA)·poly(dT) and poly d(A·T). This method measures the effect of netropsin on the thermal stability of the DNA duplex, similarly to the method used by McGhee.[57] The thermal stablity of a duplex is characterized by the temperature midpoint (t_m) seen in

the UV melting observed at 260nm. Measuring the t_m for a DNA duplex in the presence and absence of netropsin yields Δt_ms of 40°C for polyd(AT) and 45°C for poly(dA)·poly(dT). Circular dichroism was used to verifiy that there was only one type of binding in the DNA-drug complex. Binding constants were calculated using a theoretical treatment of Crothers[58]. At saturation, the ligand-induced shift in t_m is specified by:

$$\frac{1}{t_m^o} - \frac{1}{t_m} = \frac{R}{n\Delta H}\ln[1 + K(a_L)] \qquad (1.4)$$

t_m^o and t_m are the Kelvin temperatures corresponding to the midpoints of the helix-to-coil transition in the absence and presence of netropsin, respectively. ΔH is the enthalpy change for the melting of a base pair in the absence of bound ligand (from differential scanning calorimetry). K and n are the binding parameters for drug binding according to the neighbor-exclusion model (Crothers[58]) where n=1/r at saturation and K is the apparent binding constant. a_L is the free ligand activity, which equals 1/2 of the total concentration of netropsin added. These calculated binding constants refer to the association of the drug with the duplex at the high temperature of duplex melting (t_m). McGhee did not correct for this effect.[57] The binding constant can be corrected to 25°C using the relationship: $\frac{\delta \ln K}{\delta\ 1/T} = \frac{\Delta H_b}{R}$. Batch calorimetry is used to measure the ΔH_b and found that ΔC_p is 0 for both complexes. These measurements were carried out under a variety of salt conditions and are summarized in table 1.3.

Based on the slopes and intercepts of the log K *versus* log $[Na^+]$ plots, as well as comparison to observed and predicted binding constants, Breslauer *et*

Table 1.3: Equilibrium Binding Constants at 25°C for the Association of Netropsin to DNA Duplexes at Various Salt Concentrations[65]

Host Duplex	$[Na^+]$ (mM)	T_m^o (°K)	T_m (°K)	Δ (kcal/bp)	n (bp)	K (M^{-1})
poly d(AT)	16	317.9	355.9	7.3	5	2.25×10^9 M^{-1}
	66	303.4	358.9			2.13×10^8 M^{-1}
	116	335.3	358.1			8.87×10^7 M^{-1}
poly(dA)·poly(dT)	16	323.5	367.3	8.8	5	8.76×10^8 M^{-1}
	66	337.3	371.5			1.10×10^8 M^{-1}
	116	342.8	372.3			4.29×10^7 M^{-1}

al.[64] conclude that netropsin binding of the DNA involves both charged ends of the molecule. This is the observation for both poly(dA)·poly(dT), and poly d(AT).

Table 1.4: Thermodynamic Binding Profile[65]

Host Duplex	ΔG(kcal/mol)	ΔH(kcal/mol)	ΔS(cal/°K-mol)
poly d(AT)	-12.7	-11.2	+5
poly(dA)·poly(dT)	-12.2	-2.2	+33

The free energy change for binding to either poly(dA)·poly(dT) or poly d(AT) is about the same; however, the relative contributions of enthalpy and entropy are vastly different. The driving force for poly d(AT) binding is predominantly enthalpic and for poly(dA)·poly(dT) it is entropic (table 1.4). Some of the factors which could contribute to the enthalpic changes upon binding include changes in the conformation of both netropsin and the DNA as well as changes in the net hydrogen bonding in the system. One of the factors which may contribute to the entropic changes is the release of bound water. This water

could be directly displaced by the binding of netropsin to the DNA and might also result from a change in the form of the DNA resulting in an overall change in DNA hydration.

Chapter 2

Poly-*N*-methylpyrrolecarboxamide Recognition of DNA

Affinity Cleaving. Affinity cleaving is an analytical tool developed for studying DNA-small molecule interactions. It is based on the principle that attachment of a DNA cleaving functionality to a DNA binding subunit will create a bifunctional DNA binding/cleaving molecule. Restriction endonucleases and anti-tumor compounds such as bleomycin and neocarzinostatin are examples of naturally occurring molecules employing this dual binding/cleaving strategy. The first synthetic member of this class of compounds was MPE·Fe(II) prepared by Hertzberg and Dervan.[66,67] This molecule combines the known DNA intercalating functionality methidium with EDTA·Fe(II), a water-stable complex known to produce hydroxyl radicals. The bifuctional molecule, MPE·Fe(II), in the presence of molecular oxygen and a reducing agent was found to efficiently cleave double helical DNA under a variety of salt, buffer and temperature conditions.

The DNA cleaving specificity of MPE·Fe(II) was low. This result is not unexpected because the sequence specificity of the intercalator methidium is known to be low.[68] The ability of MPE·Fe(II) to recognize and cleave DNA in a sequence neutral fashion has been utilized in the technique of DNA footprinting, a method for determining the binding site of small molecules on heterogeneous DNA by cleavage inhibition.[41-43,46] Based on the findings from MPE·Fe(II), Schultz, Taylor, and Dervan explored the possibility of attaching EDTA to a sequence specific DNA binding molecule. The DNA recognition element chosen for these studies was the tri-*N*-methylpyrrolecarboxamide subunit found in the

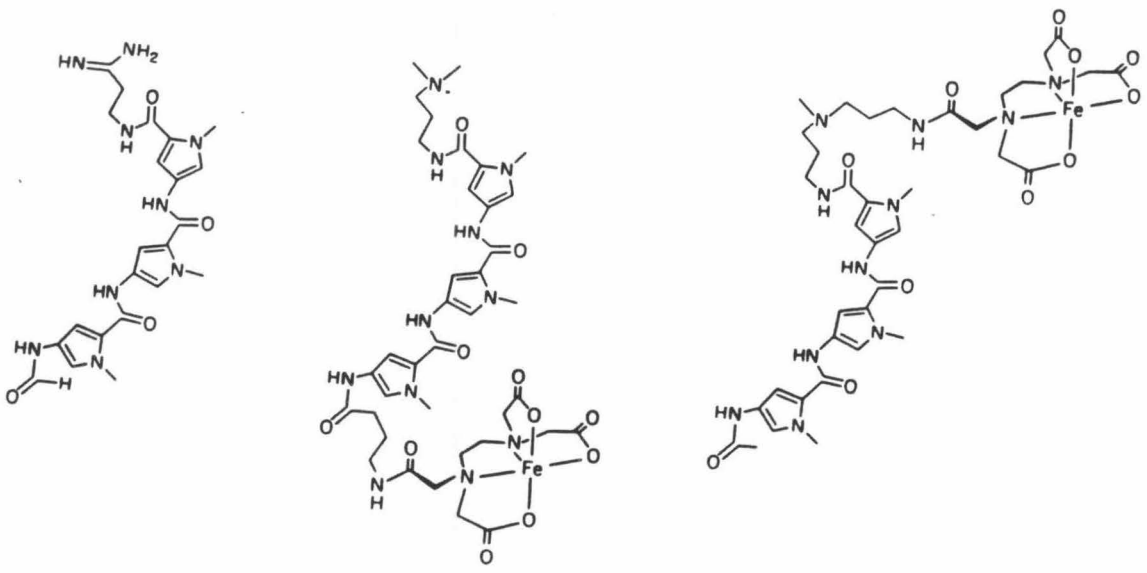


Figure 2.1 (Left) Distamycin A. (Center) DE·Fe. (Right) ED·Fe.

naturally occurring compound distamycin A, a molecule known to preferentially bind A·T rich DNA in the minor groove.[6-8, 16]

Two tri-*N*-methylpyrrolecarboxamide molecules, DE and ED, were prepared by modifying the tripeptide subunit with EDTA at either the amino- or carboxy-terminus, respectively (fig. 2.1).[69, 70] These bifunctional DNA binding/cleaving molecules, in the presence of O_2 and reducing agent, are capable of reporting their location on DNA *via* oxidative degradation of the deoxyribose backbone leading to strand scission. DNA recognition by DE and ED was studied by high resolution denaturing gel electrophoresis (fig. 2.2) on a 381 bp fragment (*Bam*H I-*Eco*R I) from pBR322 singly labeled at either the 3′ or 5′ end with ^{32}P. The cleavage reaction produced a multiple banded pattern which was quantitated

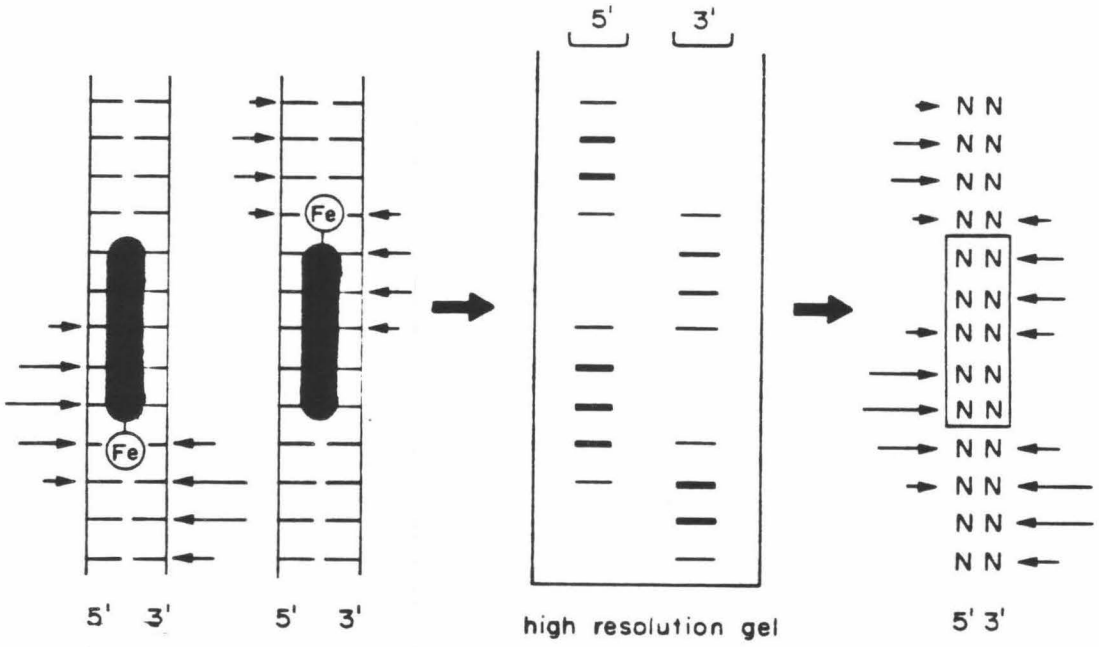


Figure 2.2 (Left) Oligopeptide-EDTA·Fe(II) complex can assume two binding orientations on the DNA affording cleavage patterns at both ends of the binding site. Electrophoresis of the reactions carried out with DNA singly labeled at either the 3′ or 5′ end with ^{32}P separates the DNA to nucleotide resolution based on charge and size (center). The DNA is visualized by autoradiography and the cleavage is quantitated by densitometry and converted into histogram form where arrow length is proportional to the amount of cleavage observed (right).

and converted into histogram form (fig. 2.3).

There are a number of lines of evidence that point to ·OH as the DNA cleaving species generated by these compounds. The DNA termini produced by MPE·Fe(II) or DE·Fe(II) cleavage of DNA are 5′ phosphate and 3′ phosphate and 3′ phosphoglycolate.[67,71] The same 3′ termini, 3′ phosphate and 3′ phosphoglycolate,[72] are observed with either EDTA·Fe(II) catalyzed cleavage or γ-radiolysis of DNA, two processes which are known to generate ·OH

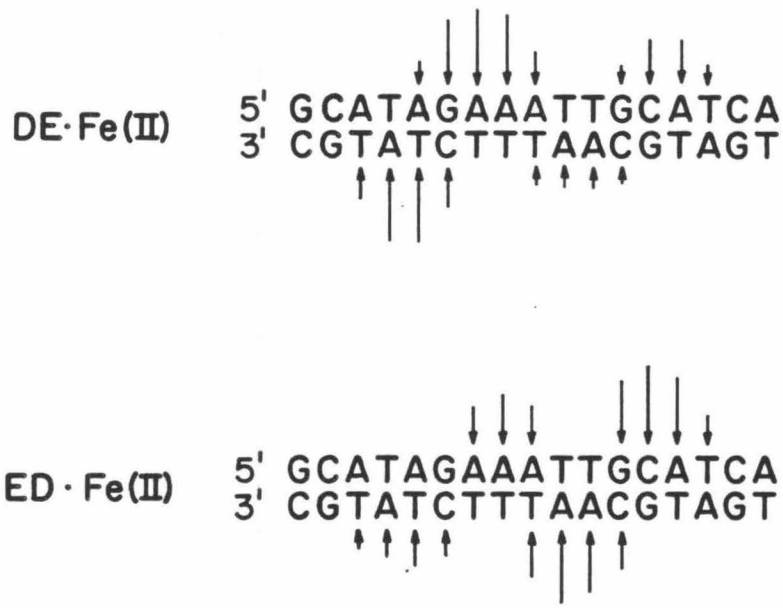


Figure 2.3 Histograms of the DNA cleavage patterns of DE·Fe(II) and ED·Fe(II) on a 381 bp restriction fragment (from ref. 70).

radicals.[66, 73] γ-radiation decomposes water as shown below:

$$H_2O \Rightarrow \cdot OH,\ e^-(\text{trapped as } {}^{-\cdot}O_2),\ \cdot H(\text{trapped as } \cdot O_2H),\ H_2$$

Udenfriend's reagent, an EDTA·Fe(II)/asorbate system, catalyzes oxidations which are consistent with the intermediacy of hydroxyl radicals.[74] Both the EDTA·Fe(II) and γ-radiation induced degradation of DNA can be inhibited by radical scavengers such as thiols or DMSO providing further evidence for a diffusible oxidizing species.[75, 76]

The multiple cut cleavage pattern seen with DE·Fe(II) also suggests a diffusible cleaving species. This pattern of cleavage is distributed over 4-5 base pairs indicating that only one face of the helix is exposed to the oxidative species as

would be expected for a diffusible species generated from a fixed position remote from the DNA. An alternate explanation for the multiple cut pattern is an iron-oxo species which is delivered to the DNA. MPE·Fe(II) has been shown to behave like a monocation in its interactions with DNA.[77] This demonstrates that the negatively charged EDTA moiety is removed from the first few shells of DNA hydration and, therefore, does not interact directly with the DNA causing it to be unable to directly deliver an oxidative species.

The specific mechanism of hydroxyl radical mediated DNA degradation is not known. Evidence from γ-radiolysis of DNA suggests that the C4′ is the hydrogen most susceptible to abstraction.[78] Once formed, this radical could then trap O_2 and form the C-4′-OOH compound. A model compound was prepared by Japanese workers to explore the mechanism of bleomycin mediated oxidative DNA strand scission (fig. 2.4).[79] Treatment of this compound with acid, even as mild as chromatography on silica gel, results in base propenal formation. Base propenal is a product from bleomycin mediated DNA degradation but it is not observed with the EDTA·Fe(II) compounds which afford only free base release.[67] Treatment of the model hydroperoxide compound under the redox active conditions of aqueous Fe(II) (0.05 M phosphate buffer, pH 7.2) and atmospheric oxygen was found to generate only free base release. More insight into the cleavage mechanism is provided by product analysis of the reaction of 2,5-dimethyl-2-hydroperoxidetetrahydrofuran and 2-hydroperoxidetetrahydrofuran with aqueous Fe(II) (fig. 2.5).[80,81] The products of this reaction are consistent

with α-bond scission as the major breakdown pathway for these hydroperoxide compounds. Based on the DNA endproduct analysis and the model compound studies, a plausible mechanism for deoxyribose degredation by ·OH is C4′ hydrogen abstraction followed by hydroperoxide formation which, under the redox conditions of the cleavage reaction, decomposes *via* α-bond scission to give free base release and DNA strand cleavage.

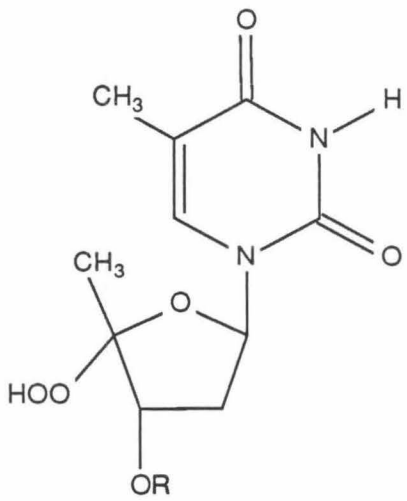


Figure 2.4 Model C-4′ Peroxide compound. R=benzyl, acetyl, or *p*-methoxybenzyl.

The multiple cut cleavage patterns produced by the DE·Fe(II) and ED·Fe(II) are shown in histogram form in figure 2.3. There are two important features to note in these patterns. The first is that changing the end of the molecule to which EDTA is attached reverses the orientation of the cleavage patterns. This result is consistent with the DNA binding orientation and specificity of the

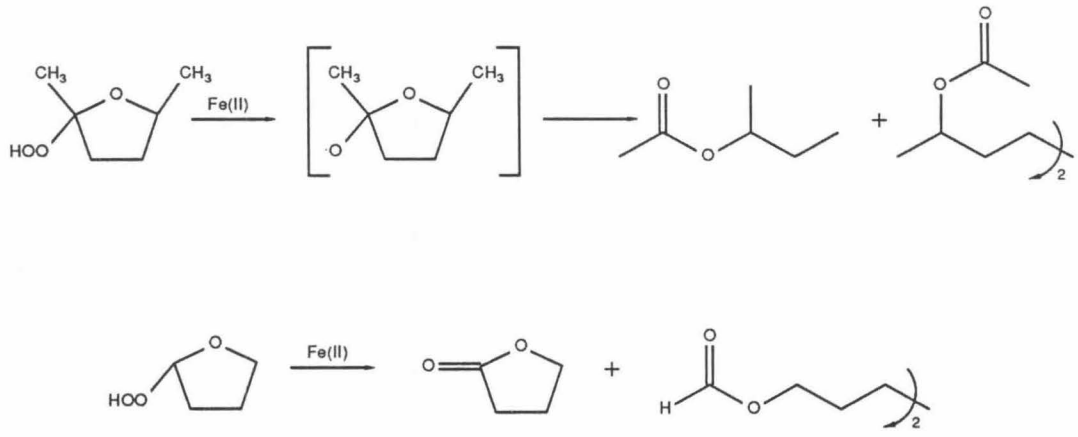


Figure 2.5

polypyrrole subunit remaining independent of EDTA attachment. The second point is that the cleavage patterns on opposite strands are shifted to the 3′ side with respect to each other. This 3′ shift results from the helical nature of the DNA (fig. 2.6). The right-handed helical twist of DNA causes deoxyribose units on opposite strands proximal to a fixed position in the minor groove to be shifted approximately two base pairs to the 3′ side in the primary structure.

The cleavage pattern can be interpreted as resulting from a diffusible cleaving species such as hydroxyl radical generated from a fixed or average position in the minor groove. The position of the EDTA·Fe(II) moiety is then assigned from the approximate two-fold rotational symmetry of the cleavage pattern. With the EDTA·Fe(II) position and knowledge of the length of the linking element, the first base pair boundary for the polypyrrole compound can be assigned (fig. 2.7). The two sets of cleavage patterns are assumed to result from two binding ori-

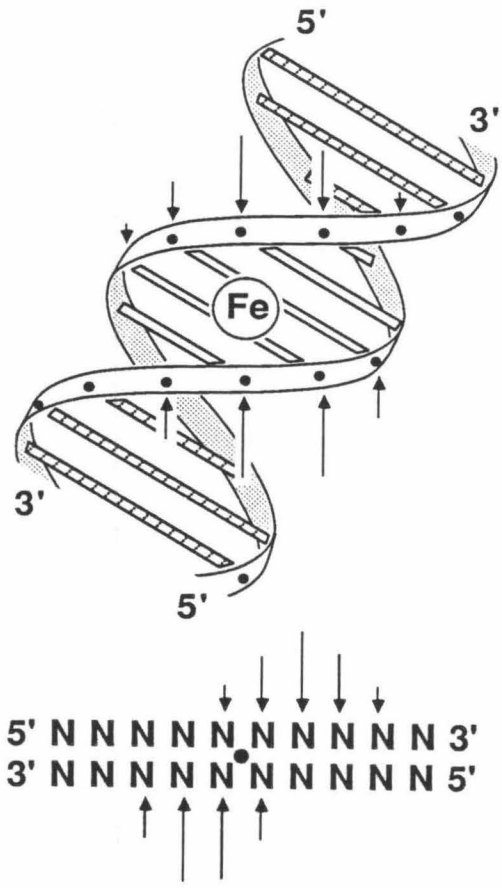


Figure 2.6 Model for Asymmetric Cleavage Pattern.

entations at the same binding site. This assumption is in accord with cleavage protection data from footprinting experiments.

Using the asymmetric cleavge model for binding site assignment and data from other labeled DNA fragments, DE·Fe(II) and ED·Fe(II) were both found to recognize the same five base pair binding sites.[69-71] This verified that the attachment of EDTA does not alter the binding specificity of the *N*-methylpyrrolecarboxamide subunit. A model for DNA recognition incorporating

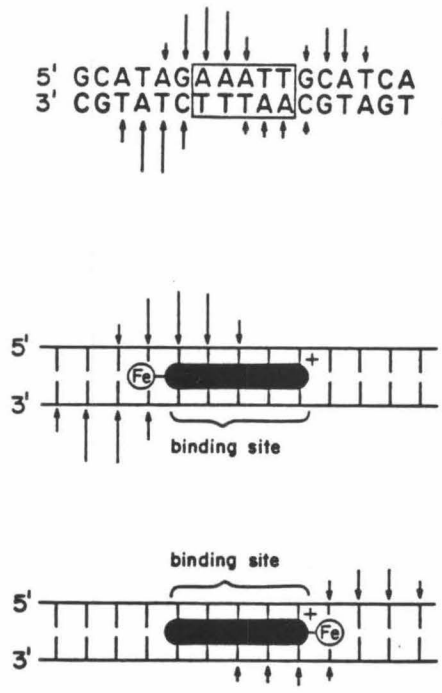


Figure 2.7 Two Binding Orientations and Assigned Site.

the binding site size data is shown in figure 2.8. It depicts the amide hydrogens forming three-center, bifurcated hydrogen bonds between adjacent A·T base pairs on opposite strands of the DNA helix. The hydrogen bonds bridge the thymine O2 and adenine N3. According to this model, the A·T specificity of these compounds results not only from the hydrogen bonds which can form with A·T base pairs but the steric obstruction in the minor groove presented by the guanine 2-amino group. A recent crystal structure of a netropsin-DNA complex has verified this as the mechanism of *N*-methylpyrrolecarboxamide recognition of DNA.[36-38]

Several groups have shown that analogs of distamycin containing four and five pyrroles bind to DNA more tightly than distamycin itself.[82-86] To study the

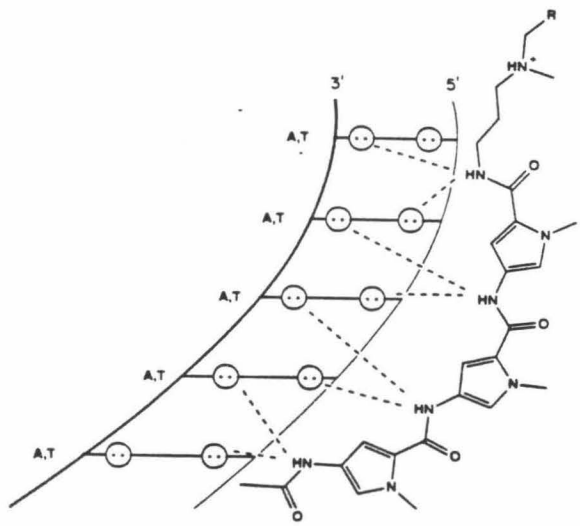


Figure 2.8 Model for poly-*N*-methylpyrrolecarboxamide recognition of DNA. Circles with two dots represent lone pairs of electrons on the N-3 of adenine or the O-2 of thymine. Dotted lines are bridged hydrogen bonds to amide NH's.

binding of an extended pyrrole chain to DNA, a five pyrrole analog of DE·Fe(II), P5E·Fe(II), was synthesized by Schultz and Dervan.[87,88] This molecule showed a larger DNA binding site size of 6-7 base pairs as compared with only 5 bp for DE·Fe(II). In addition, P5E·Fe(II) showed an increase in DNA cleavage efficiency over DE·Fe(II).

From the netropsin-DNA co-crystal, Dickerson *et al.* have suggested that the amide to amide distance in poly-*N*-methylpyrrolecarboxamides is approximately 0.1 nm too long for the base pair repeat of B-form DNA.[89] This would mean that the polypyrrole compounds should eventually fall out of register with the DNA base pairs as the pyrrole chain is extended.

The studies with DE·Fe(II), ED·Fe(II) and P5E·Fe(II) demonstrate the utility of affinity cleaving and pose number of questions. What are the factors which determine the DNA binding specificity of poly-*N*-methylpyrrolecarboxamides? How do additional pyrrole units affect a compound's binding site size and specificity and at what point do they cease to contribute to binding? To address these issues, a homologous series of poly-*N*-methylpyrrolecarboxamide-EDTA compounds containing 3-9 pyrrole units was synthesized and their recognition of DNA was studied (fig. 2.9 and 2.10). In addition, another series of acetamide compounds containing 2-7 *N*-methylpyrrolecarboxamide units were synthesized for footprinting and biophysical studies (fig. 2.11).

Synthesis. The key synthetic intermediate, tripyrrole nitro ester **14**, was prepared using a procedure reported by Bialer *et al.* with the following improvements.[90] The 4-nitro-*N*-methylpyrrolecarboxylic acid intermediate **15** can be purified from the nitration reaction mixture and the 5-nitro isomer by crystallization from 1:1 acetic acid:water. This crystallization eliminates a chromatography step and an esterification/saponification sequence. The second major improvement in the synthesis is to replace the aqueous sodium bicarbonate base in the coupling reactions with triethyl amine. These two improvements increase the overall yield of tripyrrole nitro ester **14** from ~12% to ~50% based on the commercially available *N*-methylpyrrole-2-carboxylic acid.

By procedures analogous to those used to synthesize the tri-*N*-methylpyrrole nitro ester **14**, the *N*-methylpyrrolecarboxamide chain was extended to four

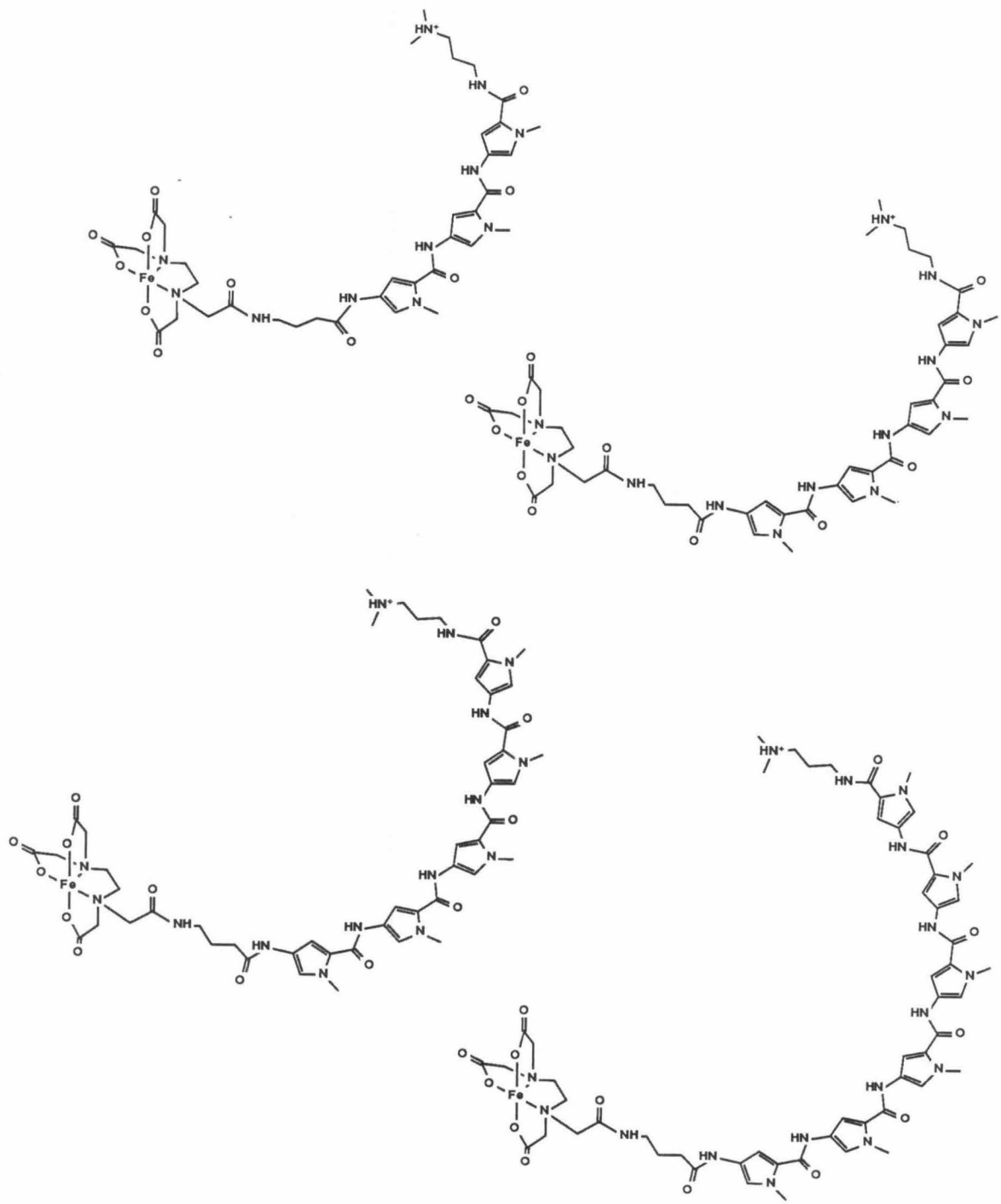


Figure 2.9 (upper left) DE·Fe (**1**·Fe), (upper right) P4E·Fe (**2**·Fe), (lower left) P5E·Fe (**3**·Fe), (lower right) P6E·Fe (**4**·Fe).

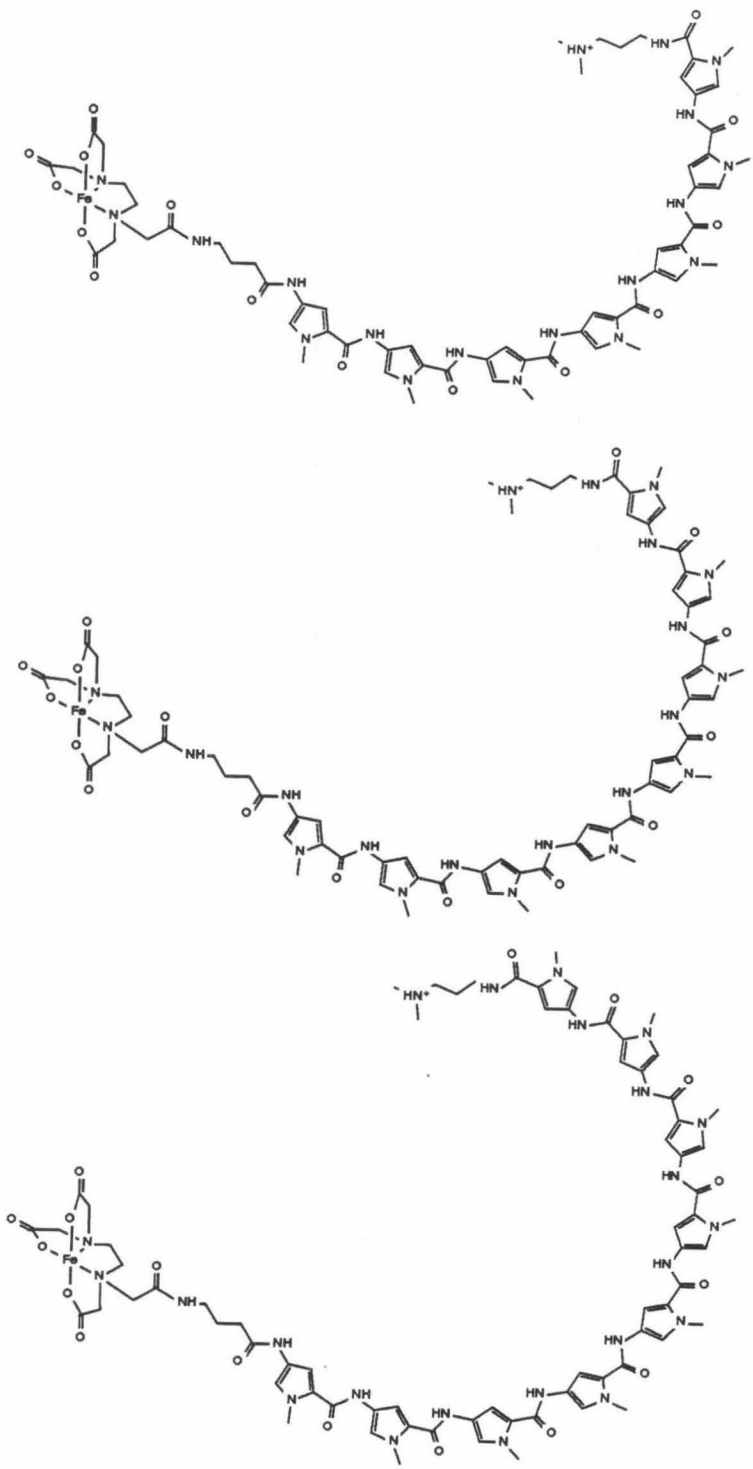


Figure 2.10 (top) P7E·Fe (**5**·Fe), (center) P8E·Fe (**6**·Fe), (lower) P9E·Fe (**7**·Fe).

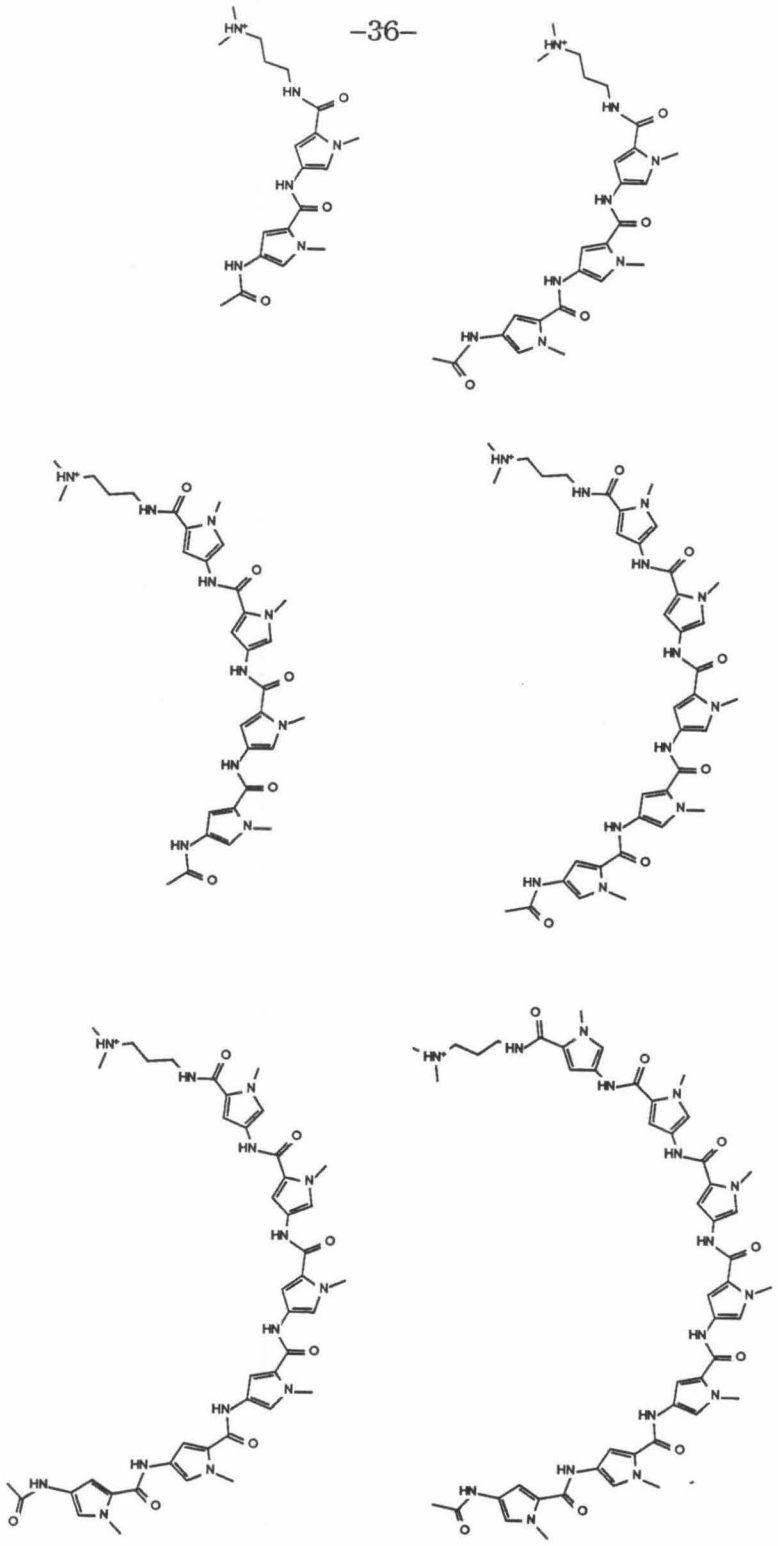
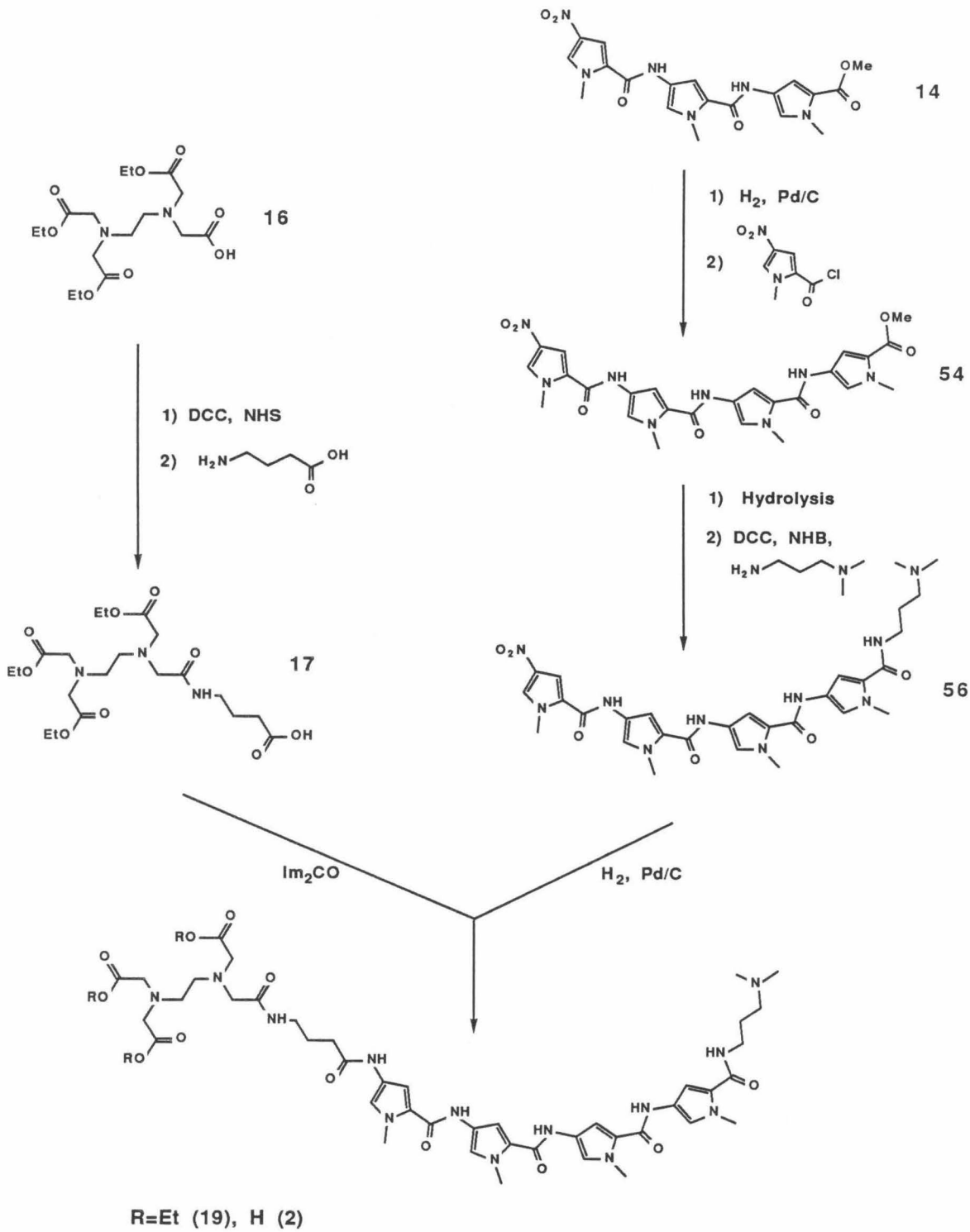


Figure 2.11 (top left) P2 (**8**), (top right) P3 (**9**), (center left) P4 (**10**), (center right) P5 (**11**), (lower left) P6 (**12**), (lower right) P7 (**13**).

and five pyrrole units (scheme 2.1). Each oligopeptide nitro ester was then hydrolyzed to the nitro acid and condensed with 3-(dimethylamino)-propylamine to give the corresponding nitro amine. Condensation of the *N*-hydroxysuccinamide ester of EDTA triethylester **16** prepared according to the method of Hay and Nolan[91] with γ-amino butyric acid produced the EDTA-triethylester-linker **17**.[70] Reduction of the corresponding nitro amines and condensation with the imidazolide of EDTA-triethylester-linker **17** produced the triesters of DE **18**, P4E **19**, and P5E **20**. The triesters were hydrolyzed with lithium hydroxide to afford the tri-, tetra-, and penta-*N*-methylpyrrolecarboxamide-EDTA molecules.

As the pyrrole chain is lengthened both the solubility of the compound and the reactivity of the terminal nitro group to reduction decrease. These two factors make the synthetic approach used for DE, P4E, and P5E untenable for the longer pyrrole compounds. The difficulties were circumvented by the synthetic route shown in scheme 2.2. The triethyl ammonium salt of the tri- or tetra-pyrrole nitro acid (**21** and **22**, respectively) was hydrogenated and coupled to the imidazolide of EDTA-triethylester-linker **17** affording the corresponding poly-*N*-methylpyrrolecarboxylic acid-EDTA-triethylester **23** or **24**, respectively. This acid was activated as the *N*-hydroxybenzotriazole ester with dicyclohexylcarbodiimide and condensed with the appropriate reduced nitro amine to give the triethylesters of P6E **25**, P7E **26**, P8E **27**, and P9E **28**. In chromatographing these triesters, it was important to use 1:1 CH_2Cl_2:MeOH with ammonium hydroxide as the solvent system to avoid precipitation of the compounds on the

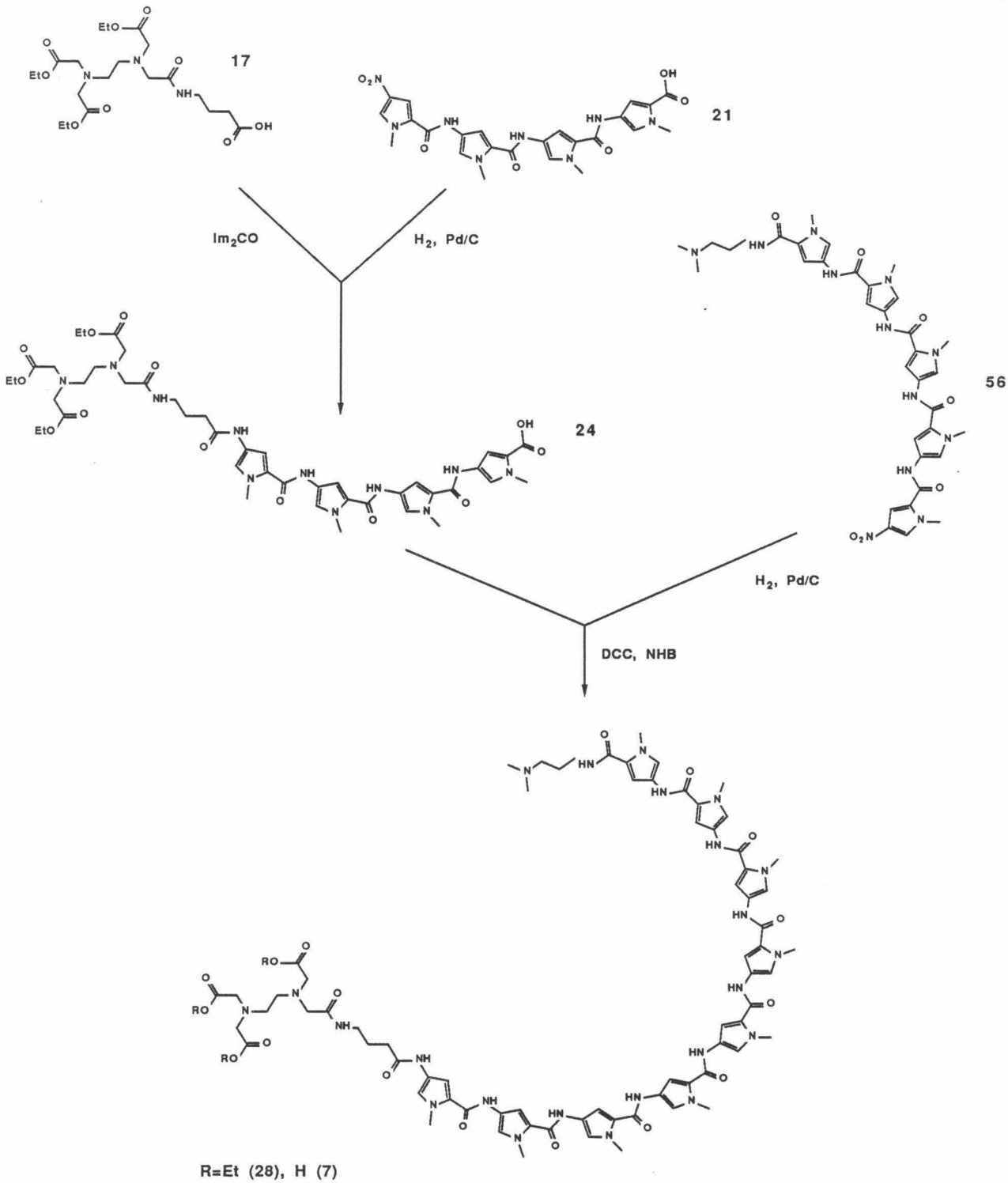


Scheme 2.1 P4E Synthesis.

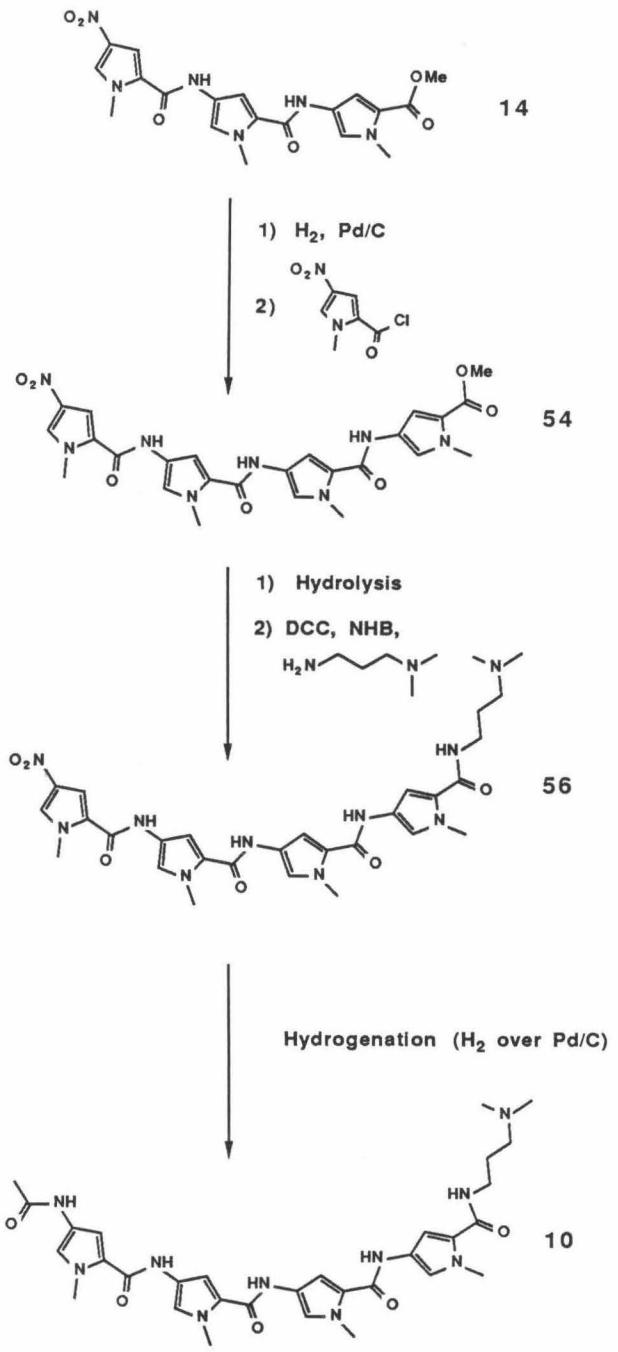
column. The triethylesters were hydrolyzed with lithium hydroxide to give the hexa-, hepta-, octa-, and nona-*N*-methylpyrrolecarboxamide-EDTA molecules.

The synthesis of the acetamide compounds was carried out similarly. The corresponding nitro amine for each of the two, three, four and five *N*-methylpyrrolecarboxamide compounds was hydrogenated at atmospheric pressure and the resulting amine was acetylated with acetyl chloride or acetyl imidazolide (scheme 2.3). Again, because of low solubility and low reactivity of the terminal nitro group to reduction, synthesis of the longer pyrrole compounds required an alternate strategy. The tripyrrole nitro ester **14** was hydrogenated under atmospheric pressure and acetylated with acetyl chloride to afford the tripyrrole acetyl ester **29**. Subsequent hydrolysis yielded the tripyrrole acetyl acid **30**. The three or four pyrrole nitro amines were hydrogenated and coupled to the tri-*N*-methylpyrrolecarboxamide acetyl acid **30** with dicyclohexylcarbodiimide and *N*-hydroxybenzotriazole to produce the hexa- and hepta-*N*-methylpyrrolecarboxamide acetyl compounds, **12** and **13**, respectively (scheme 2.4).

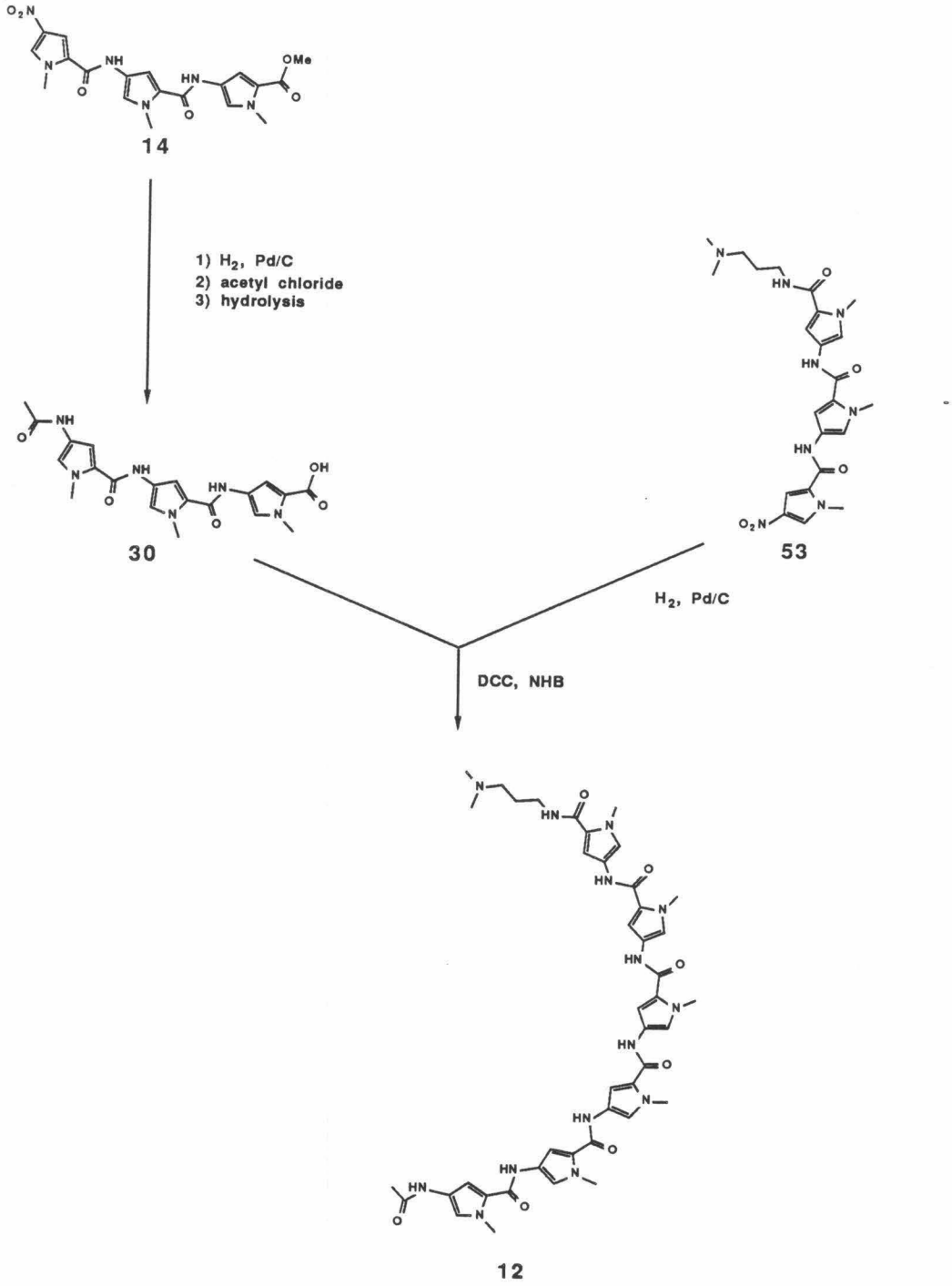
The specificities of the di-, tri-, tetra-, penta-, hexa-, and hepta-*N*-methylpyrrolecarboxamide acetamide compounds were examined by MPE·Fe(II) footprinting on a 517 bp ^{32}P-endlabeled restriction fragment from plasmid pBR322, *Eco*R I-*Rsa* I (4363-3847). The specificities of the tri-, tetra-, penta-, and hexa-*N*-methylpyrrolecarboxamide-EDTA compounds were examined on three ^{32}P-endlabeled restriction fragments from plasmid pBR322: a 167 bp fragment, *Eco*R I-*Rsa* I (1-167); a 381 bp fragment, *Bam*H I-*Eco*R I (318-1); and



Scheme 2.2 P9E Synthesis



Scheme 2.3 P4 Synthesis.



Scheme 2.4 P6 Synthesis.

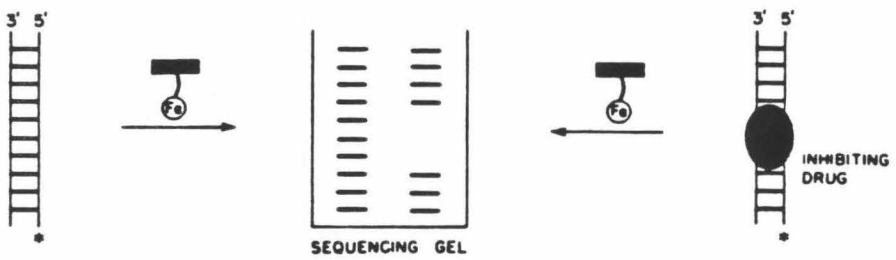


Figure 2.12 Illustration of MPE·Fe(II) footprinting on a high resolution denaturing gel with native and ligand protected segments of double helical DNA.

a 517 bp fragment, *Eco*R I-*Rsa* I (4363-3847). Because of the higher binding site requirements for the longer polypyrrole compounds, a different fragment was chosen to study the specificities of the complete series of poly-*N*-methylpyrrolecarboxamide-EDTA compounds containing 3-9 pyrroles; a 205 bp fragment of DNA containing the 5174-129 region of SV40 and 6 base pairs of the polylinker region of pUC 18 isolated from plasmid SV-CAT (*Xba* I-*Hin*d III).[92]

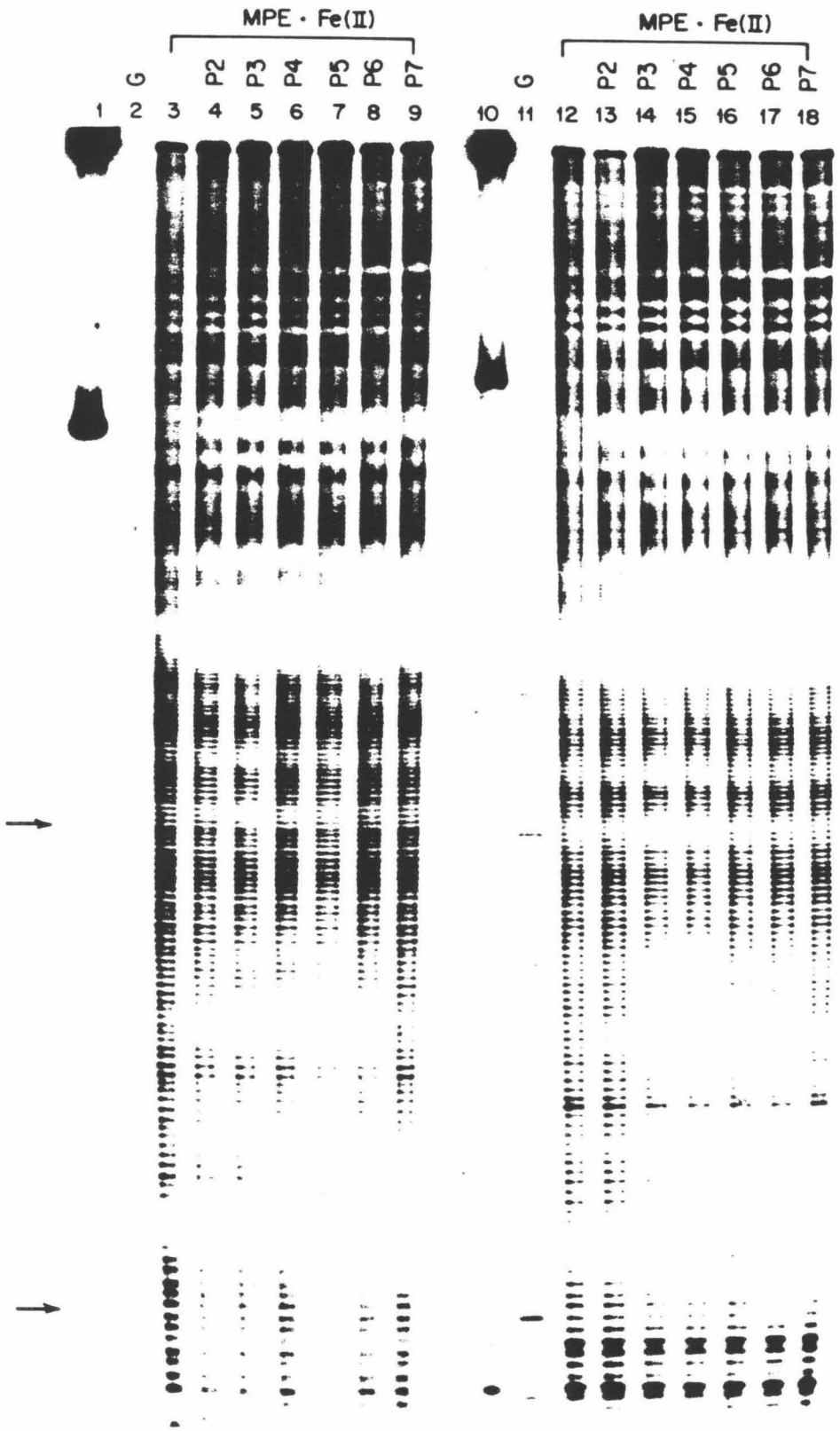
DNA Footprinting with MPE·Fe(II). DNA footprinting is a powerful direct method for determining the binding sites of proteins and small molecules on heterogeneous DNA.[41-44, 46, 93, 94] This technique is based on a sequence specific DNA binding molecule's ability to protect an endlabeled DNA fragment from cleavage at its binding site. The binding site is visualized as a region of diminished cleavage on an autoradiogram of a high resolution denaturing gel (fig. 2.12).

The footprinting reactions were run with ≥5000 cpm of a ^{32}P endlabeled 517 bp, *Eco*R I-*Rsa* I, restriction fragment from pBR322 and made up to a final concentration of 100 μM (in base pairs) with sonicated calf thymus DNA. Each of the poly-*N*-methylpyrrolecarboxamide acetamides compounds were equilibrated with the DNA at 2.5 μM (1:40 acetyl compound to bp ratio) for 1 h at 65°C. The cleavage was produced by partial digestion of the DNA with MPE·Fe(II). The reactions were carried out after equilibrating the MPE·Fe(II) at 2 μM (1:50 MPE to bp ratio) with the DNA for 15 minutes at 37°C. Cleavage was initiated by the addition of reducing agent (3 mM DTT, final concentration) and the reactions were allowed to continue for 15 minutes at 37°C. The reaction mixtures were frozen, lyophilized, and resuspended in formamide loading buffer. After heat denaturation, they were run on a 0.2-0.6 mm wedge-shaped 8% high resolution denaturing polyacrylamide gel. The gel was dried and exposed to photographic film to produce the autoradiogram shown in figure 2.13. The addition of the polypyrrole compounds produced region of diminished MPE·Fe(II) DNA cleavage. These features were quantitated by scanning the gel and plotting distance scanned versus absorbance. Binding sites were assigned from the cleavage inhibition patterns as well as from the trends in the series of compounds and visual examination of the gel. The cleavage inhibition patterns and assigned binding sites are given in histogram form in figure 2.14.

517 bp Restriction Fragment. Two common regions of cleavage inhibition are observed for the series of poly-*N*-methylpyrrolecarboxamide acetamide

Figure 2.13 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 517 bp pBR322 restriction fragment (*Eco*R I-*Rsa* I). The reactions were 100 μM in DNA base pairs (final concentration). The acetyl compounds were allowed to equilibrate with the DNA for one hour at 65°C. MPE·Fe(II) was added to a final concentration of 2 μM (1:50 MPE to bp ratio) and equilibrated with the DNA for 15 min. at 37°C. Cleavage was initiated by the addition of DTT to a final concentration of 3 mM and allowed to proceed for 15 min. at 37°C. The reaction buffer was prepared to final concentrations of 50 mM NaCl and 10 mM Tris (pH 7.4).

Lanes 1-9 are 5′-labeled DNA. Lanes 10-18 are 3′-labeled DNA. Lanes 1 and 10 are intact DNA control lanes after incubation under reaction conditions. Lanes 2 and 11 are chemical sequencing lanes containing the products of Maxam-Gilbert G-specific cleavage reactions. Lanes 3-9 and 12-18 are DNA cleavaged by MPE·Fe(II) at 2 μM. Lanes 4 and 13 contain P2 at 2.5 μM (1:50 P2 to bp ratio); Lanes 5 and 14 contain P3 at 2.5 μM (1:50 P3 to bp ratio); Lanes 6 and 15 contain P4 at 2.5 μM (1:50 P4 to bp ratio); Lanes 7 and 16 contain P5 at 2.5 μM (1:50 P5 to bp ratio); Lanes 8 and 17 contain P6 at 2.5 μM (1:50 P6 to bp ratio); Lanes 9 and 18 contain P7 at 2.5 μM (1:50 P7 to bp ratio).


MPE · Fe(II)
MPE · Fe(II)
G
G
P2
P3
P4
P5
P6
P7
1 2 3 4 5 6 7 8 9
10 11 12 13 14 15 16 17 18

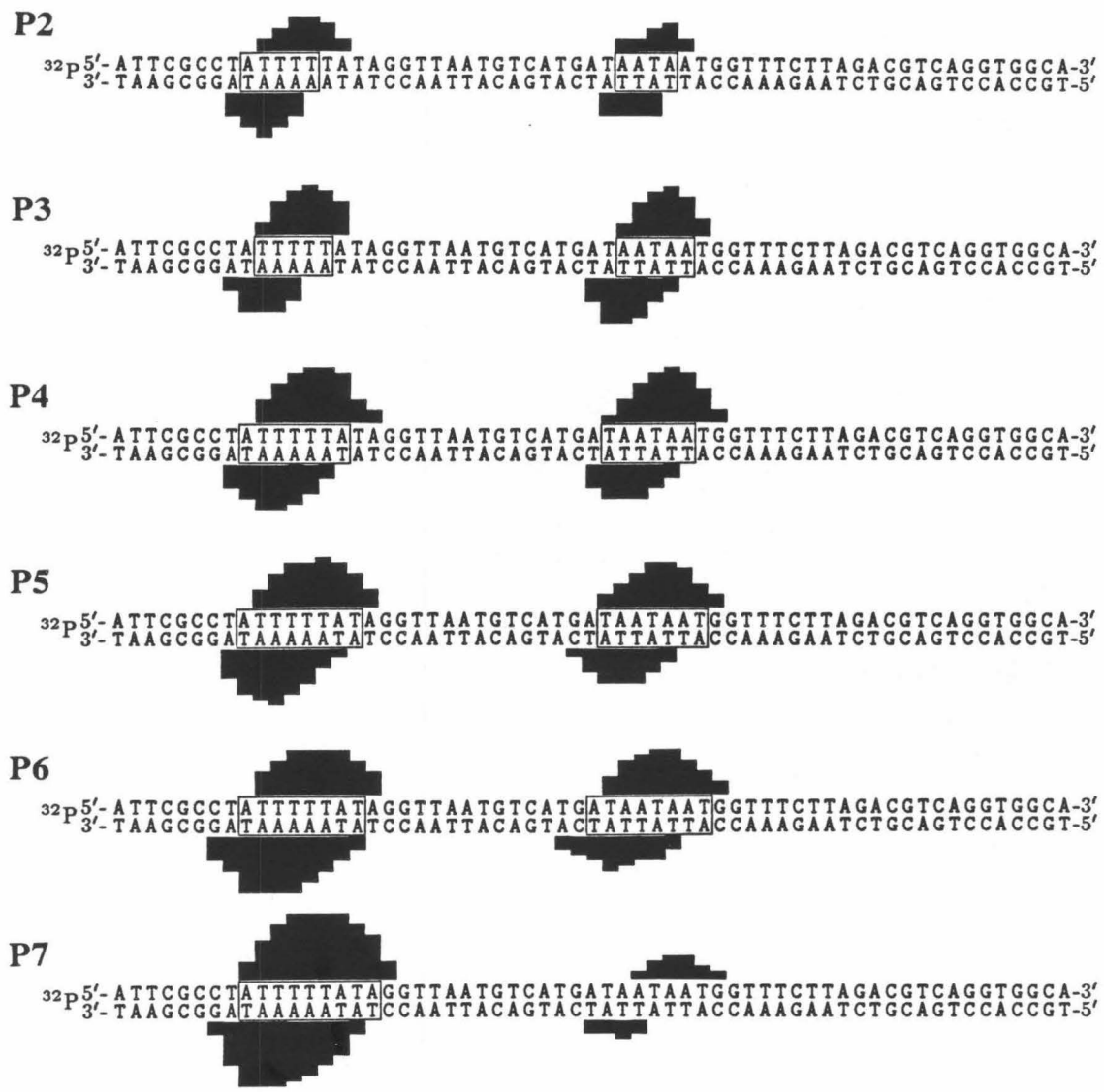


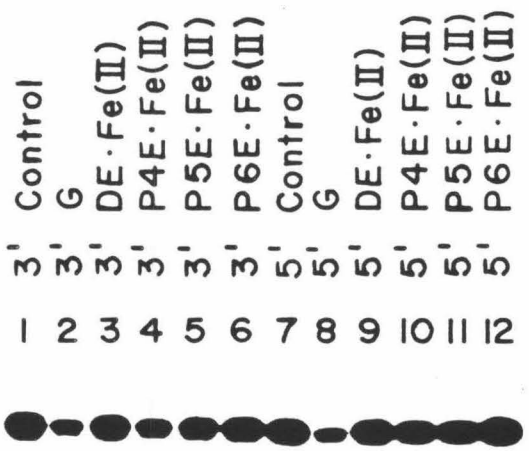
Figure 2.14 Histograms of the DNA cleavage inhibition patterns for P2, P3, P4, P5, P5, P6, and P7 on the 517 bp restriction fragment from pBR322. Bars represent the amount of MPE·Fe(II) cleavage inhibition produced by each compound. Boxes define binding site location and size.

compounds (5′–ATTTTTTATA–3′, base pairs 4331-4325 and 5′–ATAATAAT–3′, base pairs 4308-4301). The size of the cleavage inhibition patterns increases as the number of *N*-methylpyrrolecarboxamide units in the molecule increases. This increase is about one base pair per added carboxamide unit (fig. 2.14). In addition, the relative affinities of the oligopeptides for the two sites vary with P6 and P7 showing a much lower affinity for the upper site (base pairs 4308-4301).

Affinity Cleaving. The iron chelates of the polypeptide-EDTA compounds were prepared by mixing at high concentration (1-10 mM) with 1-2 equivalents of $Fe(NH_4)_2(SO_4)\cdot 6H_2O$ followed by dilution to the appropriate concentration. The cleavage reactions were carried out with $\geq$5000 cpm ^{32}P endlabeled restriction fragment and made up to a final DNA concentration of 100 μM (in base pairs) with sonicated calf thymus DNA. Each polypyrrole-EDTA·Fe(II) complex was allowed to equilibrate with the buffered DNA solution and cleavage was initiated by the addition of a dithiothreitol (DTT) solution to give a final DTT concentration of 5 mM. The final concentrations of the cleaving compounds and equilibration and reaction times and conditions are given in the gel legends. After the reactions had run to completion, the reaction mixtures were frozen, lyophilized and resuspended in formamide loading buffer. The samples were then heat denatured and run on a denaturing polyacrylamide gel. The gels were exposed to photographic film and the cleavage patterns were quantified by densitometry. The binding sites were assigned using the asymmetric cleavage model described earlier.

Figure 2.15 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 167 bp pBR322 restriction fragment (*Eco*R I-*Rsa* I). The reactions were 100 μM in DNA base pairs (final concentration). The iron chelates of the compounds were allowed to equilibrate with the DNA for one hour at 37°C. Cleavage was initiated by the addition of DTT to a final concentration of 5 mM and allowed to proceed for 2 hours at 25°C. The reaction buffer was prepared to final concentrations of 40 mM Tris base and 5 mM sodium acetate and was adjusted to pH 7.9 with acetic acid.

Lanes 1–6 are 3′-labeled DNA. Lanes 7–12 are 5′-labeled DNA. Lanes 1 and 7 are intact DNA control lanes after incubation under reaction conditions. Lanes 2 and 8 are chemical sequencing lanes containing the products of Maxam-Gilbert G-specific cleavage reactions. Lanes 3 and 9 are DNA cleavage reactions with DE·Fe(II) at 10 μM; Lanes 4 and 10 are DNA cleavage reactions with P4E·Fe(II) at 5 μM; Lanes 5 and 11 are DNA cleavage reactions with P5E·Fe(II) at 3 μM; Lanes 6 and 12 are DNA cleavage reactions with P6E·Fe(II) at 3 μM.


1 2 3 4 5 6 7 8 9 10 11 12
3' Control
3' G
3' DE·Fe(II)
3' P4E·Fe(II)
3' P5E·Fe(II)
3' P6E·Fe(II)
5' Control
5' G
5' DE·Fe(II)
5' P4E·Fe(II)
5' P5E·Fe(II)
5' P6E·Fe(II)

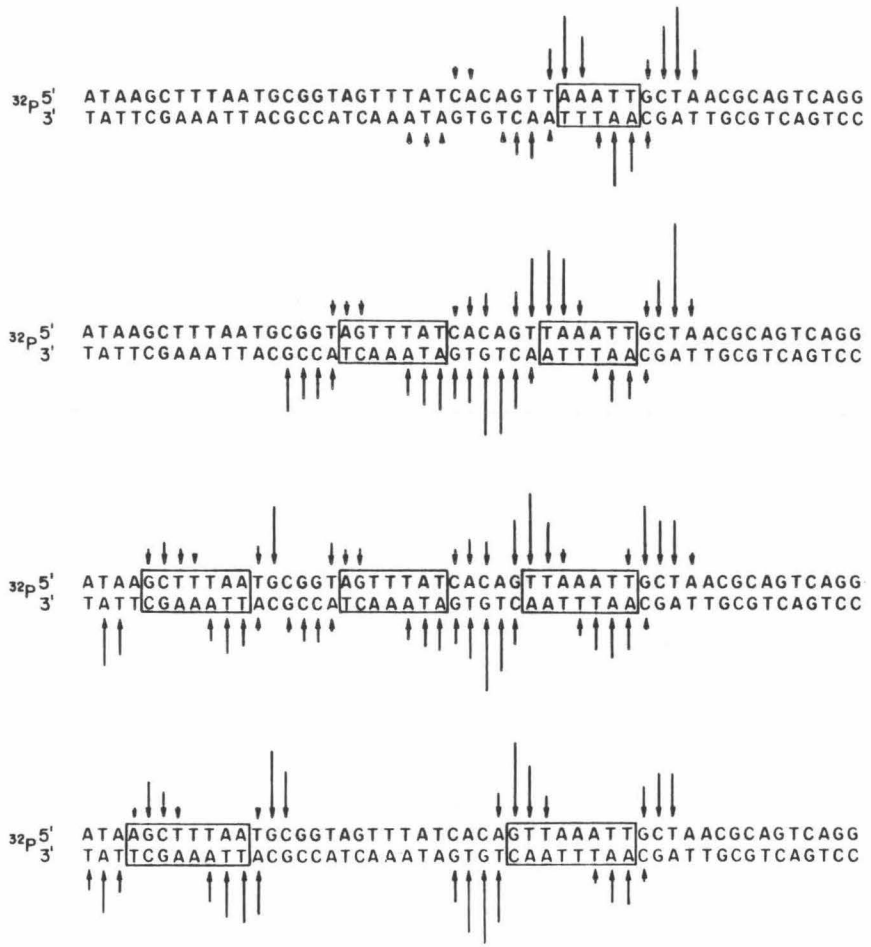


Figure 2.16 Histograms of the DNA cleavage patterns from the 167 bp restriction fragment from pBR322. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

Figure 2.17 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 381 bp pBR322 restriction fragment (*Bam*H I-*Eco*R I). The reactions were 100 μM in DNA base pairs (final concentration). The iron chelates of the compounds were allowed to equilibrate with the DNA for one hour at 37°C. Cleavage was initiated by the addition of DTT to a final concentration of 5 mM and allowed to proceed for 2 hours at 25°C. The reaction buffer was prepared to final concentrations of 40 mM Tris base and 5 mM sodium acetate and was adjusted to pH 7.9 with acetic acid.

Lanes 1–6 are 3′-labeled DNA. Lanes 7–12 are 5′-labeled DNA. Lanes 1 and 7 are intact DNA control lanes after incubation under reaction conditions. Lanes 2 and 8 are chemical sequencing lanes containing the products of Maxam-Gilbert G-specific cleavage reactions. Lanes 3 and 9 are DNA cleavage reactions with DE·Fe(II) at 10 μM; Lanes 4 and 10 are DNA cleavage reactions with P4E·Fe(II) at 5 μM; Lanes 5 and 11 are DNA cleavage reactions with P5E·Fe(II) at 3 μM; Lanes 6 and 12 are DNA cleavage reactions with P6E·Fe(II) at 3 μM.

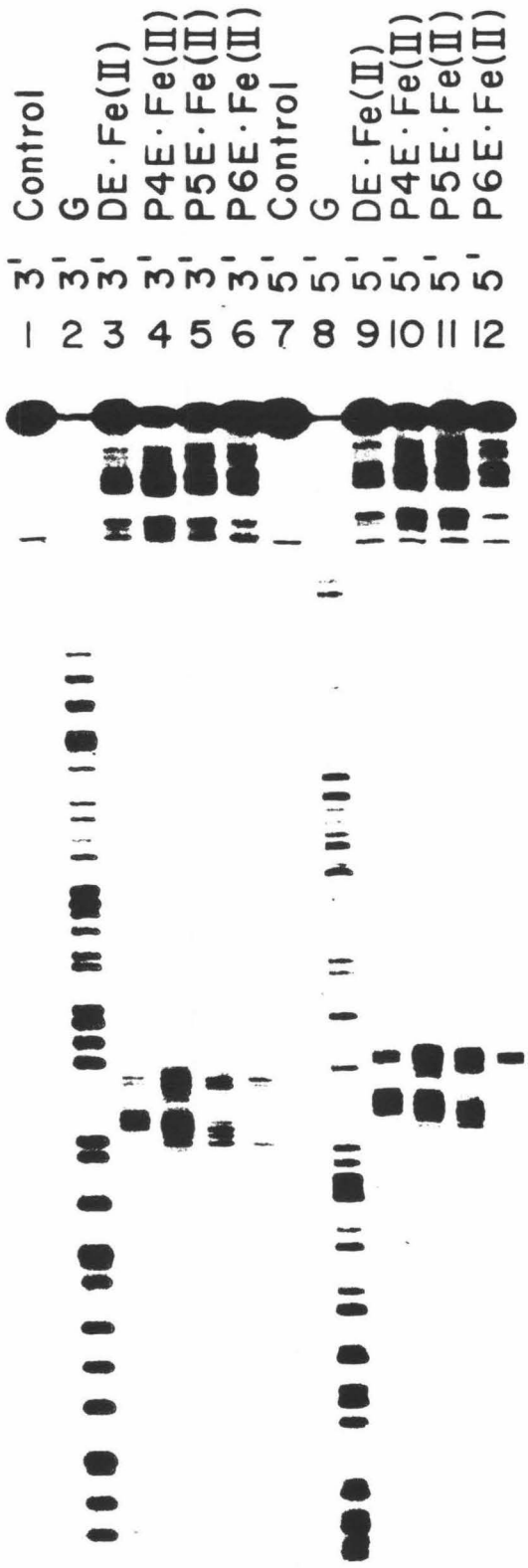

3' Control
3' G
3' DE·Fe(II)
3' P4E·Fe(II)
3' P5E·Fe(II)
3' P6E·Fe(II)
5' Control
5' G
5' DE·Fe(II)
5' P4E·Fe(II)
5' P5E·Fe(II)
5' P6E·Fe(II)
1 2 3 4 5 6 7 8 9 10 11 12

Figure 2.18 Histograms of the DNA cleavage patterns from the 381 bp restriction fragment from pBR322. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

Figure 2.19 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 517 bp pBR322 restriction fragment (*Eco*R I-*Rsa* I). The reactions were 100 μM in DNA base pairs (final concentration). The iron chelates of the compounds were allowed to equilibrate with the DNA for one hour at 37°C. Cleavage was initiated by the addition of DTT to a final concentration of 5 mM and allowed to proceed for 2 hours at 25°C.

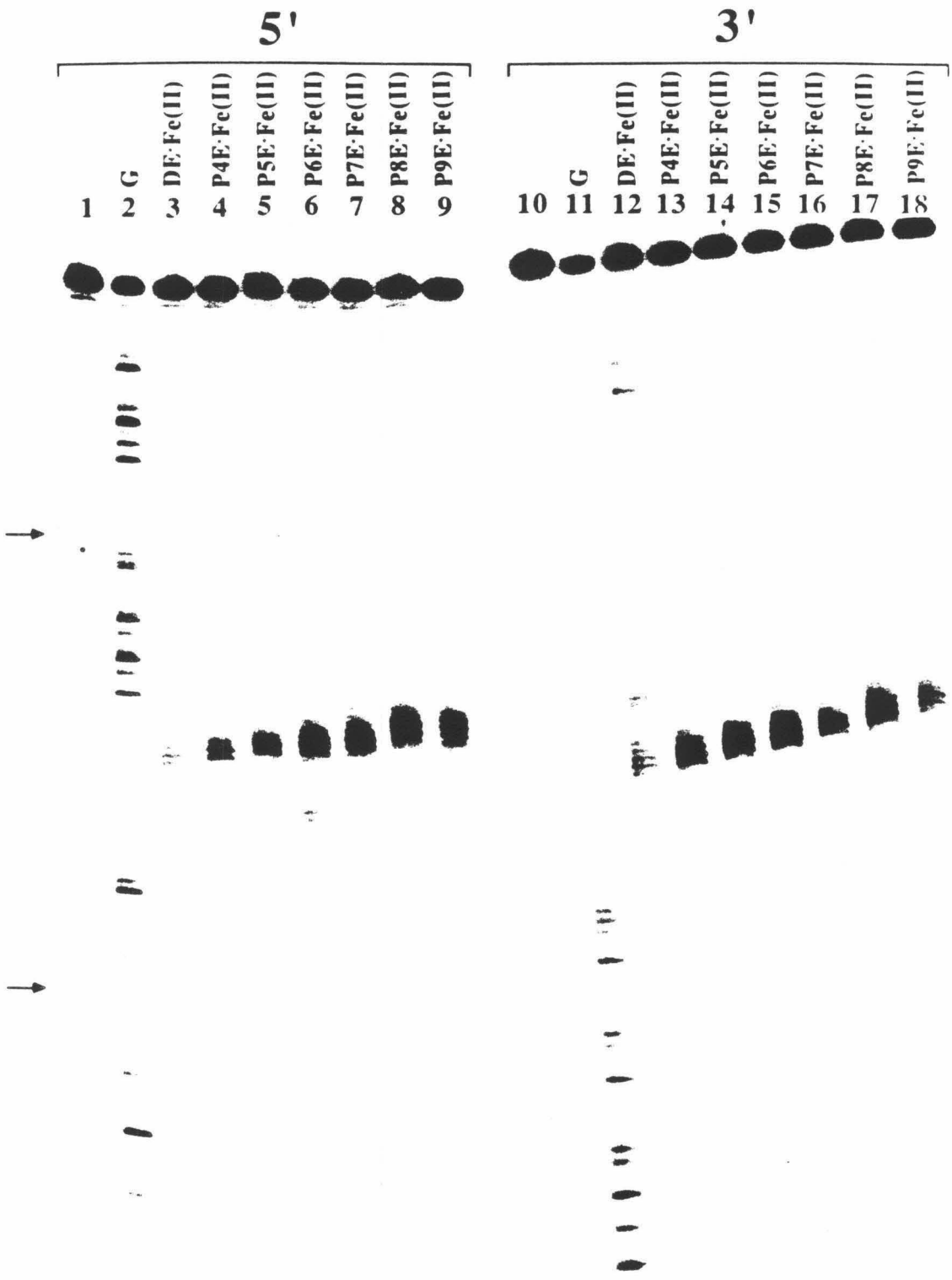
Lanes 1–6 are 3′-labeled DNA. Lanes 7–12 are 5′-labeled DNA. Lanes 1 and 7 are intact DNA control lanes after incubation under reaction conditions. Lanes 2 and 8 are chemical sequencing lanes containing the products of Maxam-Gilbert G-specific cleavage reactions. Lanes 3 and 9 are DNA cleavage reactions with DE·Fe(II) at 10 μM; Lanes 4 and 10 are DNA cleavage reactions with P4E·Fe(II) at 5 μM; Lanes 5 and 11 are DNA cleavage reactions with P5E·Fe(II) at 3 μM; Lanes 6 and 12 are DNA cleavage reactions with P6E·Fe(II) at 3 μM. The reaction buffer was prepared to final concentrations of 40 mM Tris base and 5 mM sodium acetate and was adjusted to pH 7.9 with acetic acid.


3' Control
3' G
3' DE·Fe(II)
3' P4E·Fe(II)
3' P5E·Fe(II)
3' P6E·Fe(II)
5' Control
5' G
5' DE·Fe(II)
5' P4E·Fe(II)
5' P5E·Fe(II)
5' P6E·Fe(II)
1 2 3 4 5 6 7 8 9 10 11 12

Figure 2.20 Histograms of the DNA cleavage patterns from the 517 bp restriction fragment from pBR322. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

Figure 2.21 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 205 bp restriction fragment from SV-CAT (*Xba* I-*Hin*d III). The reactions were 100 μM in DNA base pairs (final concentration). The iron chelates of the compounds were allowed to equilibrate with the DNA for one hour at 65°C. Cleavage was initiated by the addition of DTT to a final concentration of 5 mM and allowed to proceed for 2 hours at 37°C. The reaction buffer was prepared to final concentrations of 40 mM Tris base and 5 mM sodium acetate and was adjusted to pH 7.9 with acetic acid.

Lanes 1–9 are 5′-labeled DNA. Lanes 10–18 are 3′-labeled DNA. Lanes 1 and 10 are intact DNA control lanes after incubation under reaction conditions. Lanes 2 and 11 are chemical sequencing lanes containing the products of Maxam-Gilbert G-specific cleavage reactions. Lanes 3 and 12 are DNA cleavage reactions with DE·Fe(II) at 0.9 μM; Lanes 4 and 13 are DNA cleavage reactions with P4E·Fe(II) at 0.4 μM; Lanes 5 and 14 are DNA cleavage reactions with P5E·Fe(II) at 0.3 μM; Lanes 6 and 15 are DNA cleavage reactions with P6E·Fe(II) at 0.3 μM; Lanes 7 and 16 are DNA cleavage reactions with P7E·Fe(II) at 0.3 μM; Lanes 8 and 17 are DNA cleavage reactions with P8E·Fe(II) at 0.3 μM; Lanes 9 and 18 are DNA cleavage reactions with P9E·Fe(II) at 0.3 μM.


5'
3'
1
2
G
3
DE·Fe(II)
4
P4E·Fe(II)
5
P5E·Fe(II)
6
P6E·Fe(II)
7
P7E·Fe(II)
8
P8E·Fe(II)
9
P9E·Fe(II)
10
11
G
12
DE·Fe(II)
13
P4E·Fe(II)
14
P5E·Fe(II)
15
P6E·Fe(II)
16
P7E·Fe(II)
17
P8E·Fe(II)
18
P9E·Fe(II)

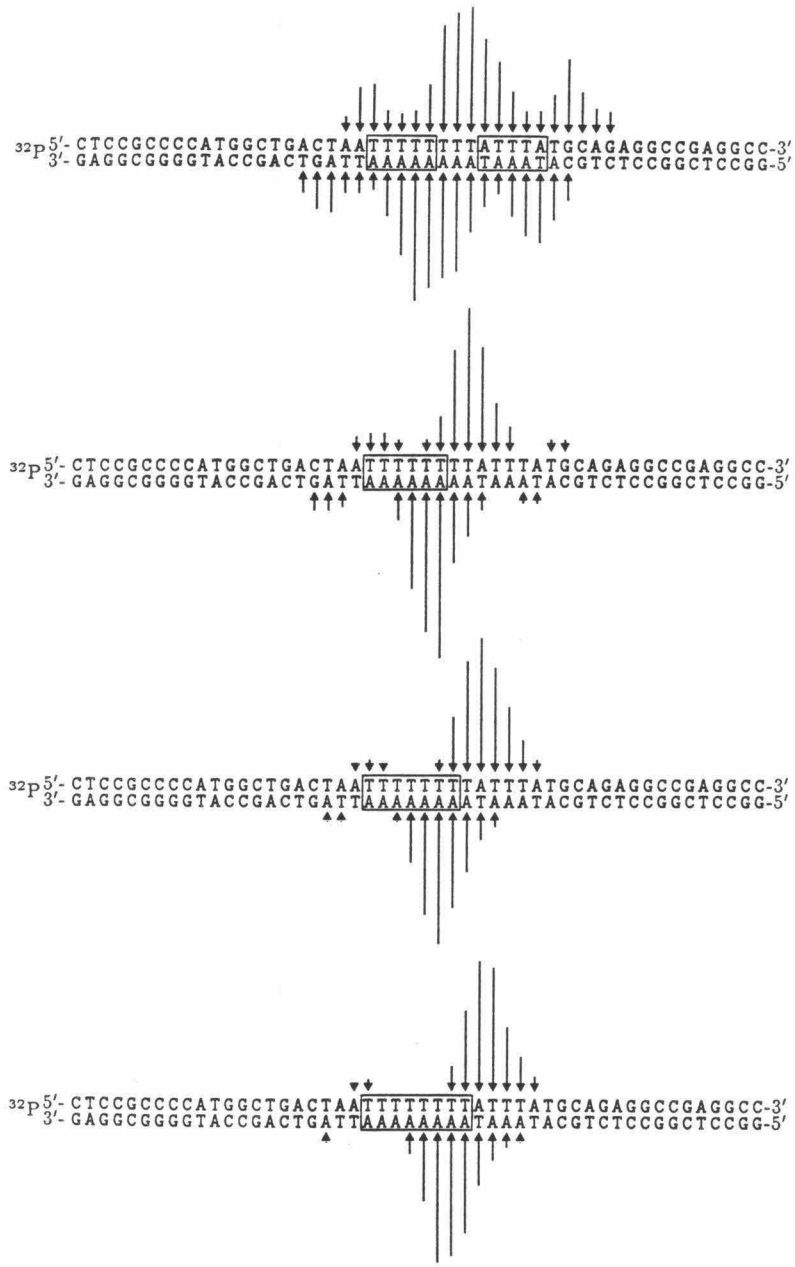


Figure 2.22 Histograms of the DNA cleavage patterns for DE·Fe(II), P4E·Fe(II), P5E·Fe(II), and P6E·Fe(II) on the 205 bp restriction fragment from SV-CAT. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

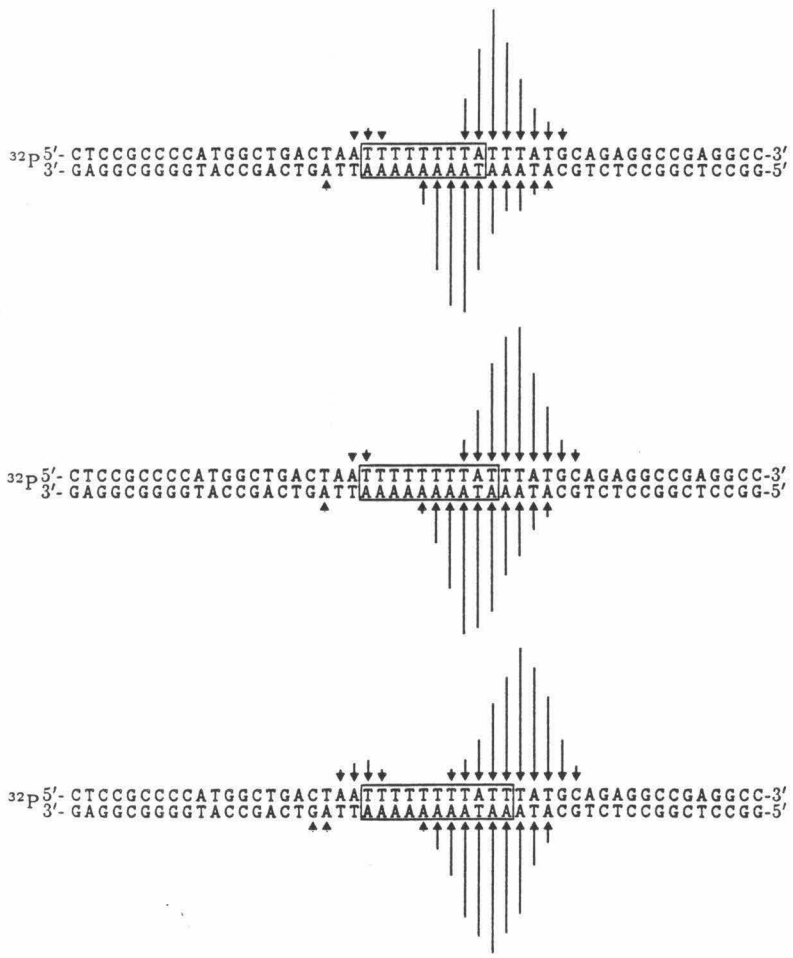


Figure 2.23 Histograms of the DNA cleavage patterns for P7E·Fe(II), P8E·Fe(II), and P9E·Fe(II) on the 205 bp restriction fragment from SV-CAT. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

167 bp Restriction Fragment. This restriction fragment shows one binding site that is conserved for all four compounds, DE·Fe(II)-P6E·Fe(II) (5′–GTTAAATT–3′, base pairs 65-72 on pBR322). This high affinity site is easily resolved on the autoradiograms and shows an increase in the recognition sequence of one base for each additional *N*-methylpyrrolecarboxamide unit added in the series. There is a slight preference for the amino terminus of the polypyrrole binding unit to orient towards the 5′ end of this site (figs. 2.15 and 2.16).

This fragment also contains minor sites with lower cleavage intensity for DE·Fe(II), P4E·Fe(II), and P5E·Fe(II) in the region of base pairs 30-38, 5′–AGTTTAT–3′, on pBR322. These sites are assigned binding sites of 5 bp for DE·Fe(II) and 7 bp for P4E·Fe(II) and P5E·Fe(II). An orientation preference is difficult to assign because of the small amount of cleavage. This site does not appear in the P6E·Fe(II) lane.

The final site on this fragment, 5′–AGCTTTAA–3′, is also weak and is only resolved for P5E·Fe(II) and P6E·Fe(II). In the case of the smaller compounds an increased cleavage background is seen in this area, but no defined binding site is apparent. With P5E·Fe(II) and P6E·Fe(II), 7 and 8 base pair binding sites, respectively, are observed.

381 bp Restriction Fragment. One common binding location is seen for each of the four homologous oligopeptide-EDTA compounds on the 381 bp restriction fragment (5′–TAGAAATT–3′, base pairs 252-259 on pBR322). At this location, there is an increase in binding site size of one base pair with each

additional *N*-methylpyrrolecarboxamide unit. In addition, the orientation preference of the oligopeptide-EDTA changes with increasing chain length. The DE·Fe(II) pattern shows a bias for the tripeptide amino terminus to orient towards the 5′ end of the binding site, 5′–AAATT–3′. The tetra- and penta-*N*-methylpyrrolecarboxamides, P4E·Fe(II) and P5E·Fe(II), show equi-energetic orientational preference. With P6E·Fe(II), the amino terminus of the oligopeptide orients toward the 3′ end of the sequence, 5′–TAGAAATT–3′ (figs. 2.17 and 2.18).

517 bp Restriction Fragment. Two common binding locations are observed for the four oligopeptide-EDTA·Fe(II) molecules. At both sites there is an increase in binding site size of one base pair for each additional *N*-methylpyrrolecarboxamide unit (figs. 2.19, 2.20). All four oligopeptides preferentially orient their amino terminus towards the 3′ side of the lower binding site (5′–ATTTTTAT–3′, base pairs 4331-4324). There is little orientation preference at the 5′–ATAATAAT–3′site (base pairs 4308-4301). In addition, the relative affinities of the oligopeptides for the two sites change as a function of the number of *N*-methylpyrrole groups. For DE and P6E the base pair 4331-4324 location appears to be the stronger of the two binding sites.

SV-CAT 205 bp Restriction Fragment. This fragment contains one common binding locus for all seven polypyrrole compounds (5′–TTTTTTTTA-TT–3′, 18-28 on SV-40). DE·Fe(II) has two well defined binding sites within this region and one poorly resolved site much farther up the gel. As the polypyrrole

chain is extended, the unresolved upper site becomes weaker until it is no longer apparent in the compounds with five or more pyrroles. At the lower, well resolved site the series of compounds shows cleavage consistent with an increase in binding site size of one base pair with each additional pyrrole subunit (figs. 2.21, 2.22 and 2.23). There is also a very strong orientation preference for the compounds to orient their amino terminus towards the 3′ end of the run of thymines.

Binding Site Size. The series of poly-*N*-methylpyrrolecarboxamide acetamide compounds inhibit MPE·Fe(II) mediated DNA degradation of the 517 bp fragment at two strong common locations. The binding site sequences and sizes for the series of acetamide compounds are given in table 2.1. The four homologous oligopeptide·Fe(II) compounds, DE·Fe(II)-P6E·Fe(II), cleave the three pBR322 restriction fragments strongly at four common locations. The series of DNA cleaving compounds containing 3-9 *N*-methylpyrrolecarboxamide units also cleave a restriction fragment containing the SV40 promoter region at a single strong common site. The binding site sizes and sequences for the affinity cleaving compounds, DE, P4E, P5E, P6E, P7E, P8E, and P9E are given in table 2.2. The footprinting results are consistent with the two cleavage patterns observed with the affinity cleaving compounds resulting from two binding orientations at a single recognition sequence. This establishes the validity of the asymmetric cleavage model used for binding site assignment with the affinity cleaving compounds. The observation of the same binding sites with both the affinity cleaving compounds and the acetamide compounds also demonstrates that the

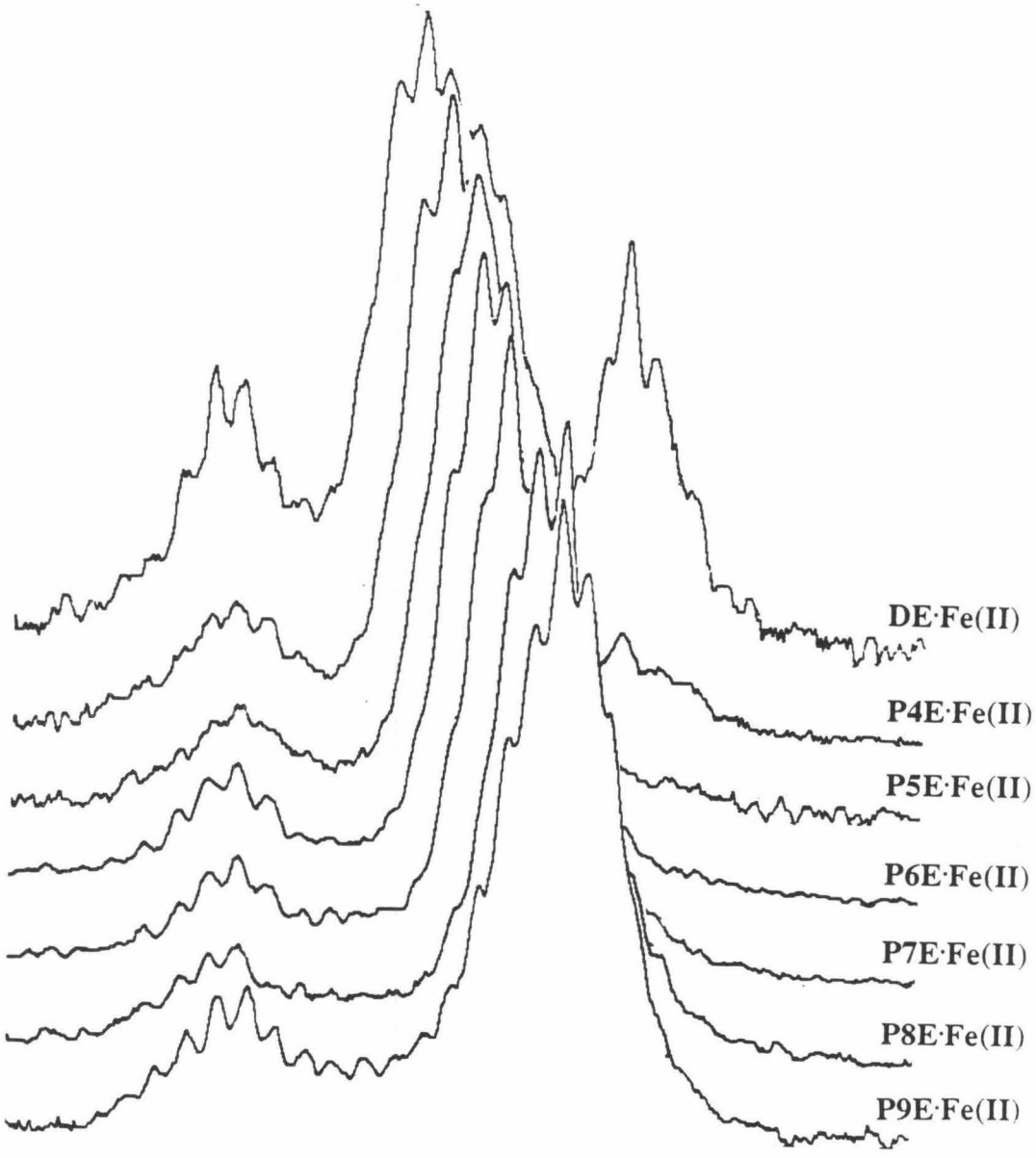


Figure 2.24 Densitometry Scans of the 5′ data from the high resolution gels of the affinity cleaving compounds on the SV40 promoter aligned by sequence (from fig. 2.21).

attachment of EDTA has no apparent effect on the binding site selection by the poly-*N*-methylpyrrolecarboxamide recognition element.

Table 2.1: Oligopeptide binding sites on the 517 restriction fragment observed by MPE·Fe(II) footprinting

Oligo-peptide	n*	Site(5′-3′)	Site size (base pairs)
P2	3	A T T T T t a t a	5
P3	4	a T T T T T a t a	5
P4	5	A T T T T T A t a	7
P5	6	A T T T T T A T a	8
P6	7	A T T T T T A T a	8
P7	8	A T T T T T A T A	9
P2	3	a t A A T A a t	4
P3	4	a t A A T A A t	5
P4	5	a T A A T A A t	6
P5	6	a T A A T A A T	7
P6	7	A T A A T A A T	8

*Number of amide units.

In each sequence, bases contained in the site are symbolized with capital letters and the the neighboring nucleotides are designated by lower case letters.

The oligopeptides with three, four, five, six, seven, eight and nine *N*-methylpyrrolecarboxamide units containing four, five, six, seven, eight, nine, and ten amide NH's, bind sites of A·T-rich DNA consisting of five, six, seven, eight, nine, ten, and eleven contiguous base pairs, respectively. This result leads to a general rule for poly-*N*-methylpyrrolecarboxamide recognition of DNA in which a molecule containing n amides recognizes a binding site of $n + 1$ base pairs. This rule is consistent with the oligopeptides binding in the minor groove of right-handed DNA with the amide NH groups forming three center bridging hydrogen bonds between adjacent N-3 and O-2 atoms of adenine or thymine, respectively, on opposite strands of the DNA helix (fig. 2.8).

Table 2.2: Oligopeptide binding sites on the 517, 167, 381, and SV40 restriction fragments observed by affinity cleaving

Oligo-peptide	n*	Site(5′-3′)	Site size (base pairs)
		517 bp Fragment	
DE	4	a T T T T T a t	5
P4E	5	a T T T T T A t	6
P5E	6	a T T T T T A T	7
P6E	7	A T T T T T A T	8
DE	4	a t A A T A A t	5
P4E	5	a T A A T A A t	6
P5E	6	a T A A T A A T	7
P6E	7	A T A A T A A T	8
		167 bp Fragment	
DE	4	g t t A A A T T	5
P4E	5	g t T A A A T T	6
P5E	6	g T T A A A T T	7
P6E	7	G T T A A A T T	8
		381 bp Fragment	
DE	4	t a g A A A T T	5
P4E	5	t a G A A A T T	6
P5E	6	t A G A A A T T	7
P6E	7	T A G A A A T T	8
		SV40 Fragment	
DE	4	T T T T T t t t a t t	5
P4E	5	T T T T T T t t a t t	6
P5E	6	T T T T T T T t a t t	7
P6E	7	T T T T T T T T a t t	8
P7E	8	T T T T T T T T A t t	9
P8E	9	T T T T T T T T A T t	10
P9E	10	T T T T T T T T A T T	11

*Number of amide units.

In each sequence, bases contained in the site are symbolized with capital letters and the the neighboring nucleotides are designated by lower case letters.

The largest of these compounds, P9E·Fe(II), contains 10 amides and recognizes 11 bp of DNA. Distamycin and netropsin are known to bind B-form DNA[17–22] and the helical repeat of B-form DNA is 10.4 bp/turn.[95] Therefore, if the DNA-P9E·Fe(II) complex is in the B-form, P9E·Fe(II) is recognizing more than a complete turn of the DNA helix.

Orientation. From the relative intensity of the two affinity cleavage patterns flanking the binding sites, one can estimate the relative orientation preference of the oligopeptide on DNA. A number of factors may influence the orientation of the pyrrole compound at its binding site including: the local sequence, the flanking sequence and the number of *N*-methylpyrrolecarboxamides in the molecule. It is important to note that, the two faces of the pyrrole ring are prochiral; therefore, they have a diastereomeric interaction with the chiral DNA which could potentially result in a difference in energy for the two binding orientations. The binding orientations could also be effected by the molecule interacting preferentially with one or the other of the two strands of DNA at the binding site. This interpretation is not supported by the observations from the co-crystal of netropsin and DNA which indicate that the polypyrrole molecule binds in the center of the minor groove. However, this crystal structure showed a significantly shorter distance between the amide NH's and the thymine O2's than with the adenine N3's (table 2.3). Additionally, netropsin may sit in the center of the minor groove of the dodecamer because the sequence is symmetrical. There is also evidence that the minor groove in the crystal structure may

be anomalously narrow.[39,40] On an unsymmetrical sequence, the minor groove may be wider allowing the poly-*N*-methylpyrrolecarboxamides to preferentially associate with one of the strands possibly the thymine rich strand. Because the edge of the pyrrole interacting with the minor groove is not symmetrical, one binding orientation should provide better H-bonding than the other. Sites containing long runs of poly dA·poly dT show a very pronounced preference for the amino terminus of the polypyrrole subunit to orient towards at the 3′ end of the pyrimidine strand (figs. 2.22 and 2.23). This suggests that, if hydrogen bond formation is the controlling factor in binding orientation, this orientation provides optimal interaction between the amide NH's and the hydrogen bond acceptors on poly(dA)·poly(dT) possibly the thymine O2 in particular.

Table 2.3: Crystal Structure H-bond Distances[37]

Netropsin atom	DNA atom	Distance Å	Δ Distances Å
N-4	O-2 of T-20	2.68	0.60
N-4	N-3 of A-18	3.28	
N-6	O-2 of T-19	3.86	0.34
N-6	O-2 of T-7	3.54	
N-8	N-3 of A-18	3.45	0.89
N-8	O-2 of T-8	2.56	

Atom numbers refer those in figure 1.2 from ref. 37

The influence of flanking sequences on the local structure of the binding site may be important in DNA recognition. An example of the effect of flanking sequences can be seen in a comparison of two DE·Fe(II) sites. On the 381

bp fragment, the sequence 5′–agAAATTgc–3′ has a 70:30 orientation preference for the amino end of the tripyrrole subunit of DE to lie at the 5′ side of the site. The 167 bp fragment with DE shows *no* orientation preference at the site 5′–ttAAATTgc–3′. Since the binding sites are identical except they occur in different contexts, the difference in orientation at the recognition sequence must be induced by the flanking sequences and the alterations they impose on the local structure. One mechanism by which neighboring regions could exert their effect on binding sequences is by influencing the minor groove width. Distamycin footprinting studies with DNase have been reported to show cleavage modulations consistent with changes in the minor groove width.[93] This demonstrates that there is a relationship between poly-*N*-methylpyrrolecarboxamides binding to DNA and this structural variable. Because of their sensitivity to DNA tertiary structure, these oligopeptide-EDTA molecules may be useful as sensitive probes of variations in B-form DNA and the influence of flanking sequences.

Salt Effect. Because high salt appears to break up weak and non-specific interactions between poly-*N*-methylpyrrolecarboxamides and DNA,[7,8] P6E·Fe(II) cleavage reactions were run with a range of 0-2 M added NaCl (fig. 2.25). Two major effects of increasing salt concentration were observed. The first observation was that some binding sites faded out as the salt concentration increased. The second was that the cleavage patterns became broader covering more bases.

Increasing the salt concentration in affinity cleaving reactions has two major consequences: it decreases the strength of electrostatic interactions and the

Figure 2.25 The effect of added salt on the affinity cleaving reaction. Autoradiograms of high resolution denaturing gel electrophoresis pBR322 restriction fragments. (Left) 5′ endlabeled 167 bp fragment (*Eco*R I-*Rsa* I) and (Right) 5′ endlabeled 517 bp fragment (*Eco*R I-*Rsa* I). The reactions were 100 μM in DNA base pairs final concentration. The P6E·Fe(II) was allowed to equilibrate with the DNA for 1 hour at 37°C. Cleavage was initiated by the addition of DTT to a final concentration of 5 mM and allowed to proceed for 2 hours at 25°C. Reactions were halted by ethanol precipitation and were washed once with 70% ethanol/water before being taken up in loading buffer. The reaction buffer was prepared to final concentrations of 40 mM Tris base and 5 mM sodium acetate and was adjusted to pH 7.9 with acetic acid.

Lanes 1 and 9 are intact DNA controls after incubation under reaction conditions. Lane 2 and 10 are chemical sequencing lanes containing the products of of Maxam-Gilbert G-specific cleavage reactions. All reactions in lanes 3–8 and 11–16 were carried out with P6E·Fe(II) at 3 μM. Lanes 3 and 11 contain no added sodium chloride; Lanes 4 and 12 contain 0.1 M added sodium chloride; Lanes 5 and 13 contain 0.25 M added sodium chloride; Lanes 6 and 14 contain 0.5 M added sodium chloride; Lanes 7 and 15 contain 1 M added sodium chloride; Lanes 8 and 16 contain 2 M added sodium chloride.


1 2 3 4 5 6 7 8
1 2 3 4 5 6 7 8

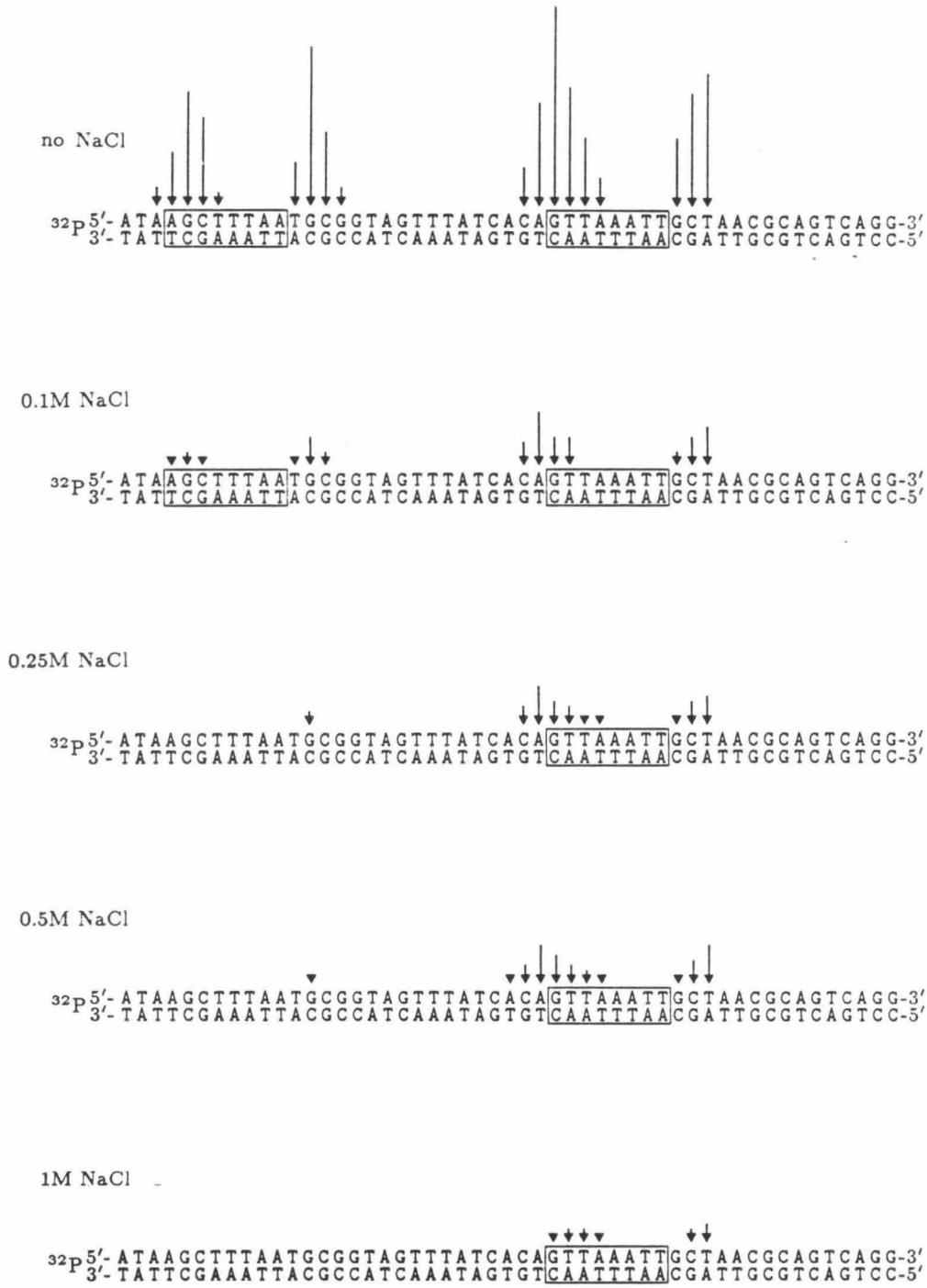


Figure 2.26 Histograms of the DNA cleavage patterns from cleavage of the 167 bp restriction fragment shown in figure 2.25. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

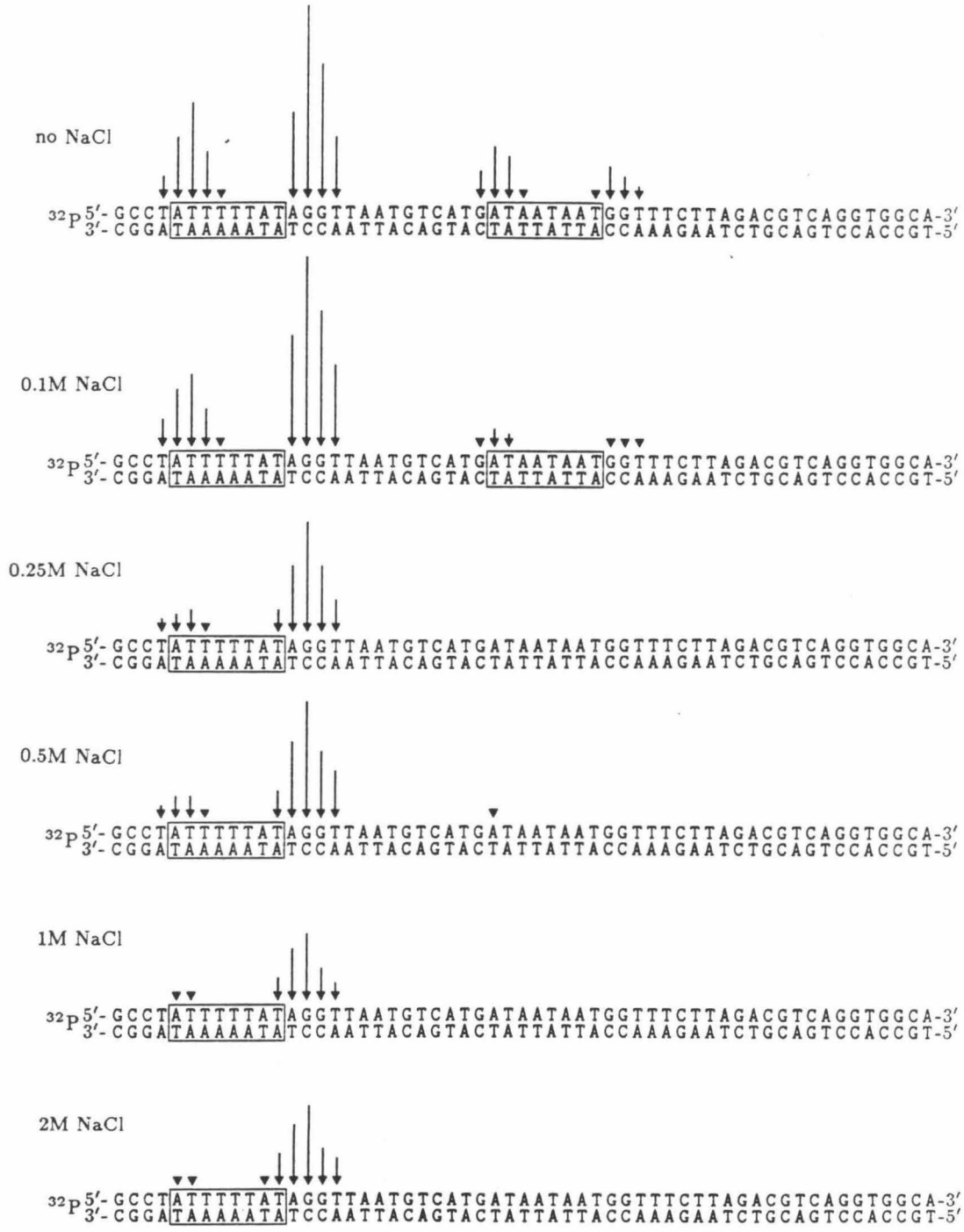


Figure 2.27 Histograms of the DNA cleavage patterns from cleavage of the 517 bp restriction fragment from shown in figure 2.25. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

salt cations compete with $Fe^{2+/3+}$ for the EDTA. These two factors combine to decrease the amount of cleavage observed with affinity cleaving compounds by making the DNA cleaving functionality less efficient and by lowering the binding constant of the binding portion of the molecule as a result of lowering the attraction between the positively charged amino terminus of the molecule and the negatively charged DNA.

Interestingly, the cleavage does not decline uniformly at all of the sites. On the 167 bp fragment one of the sites is cleaved at 1 M added NaCl and with the 517 bp fragment efficient cleavage is still observed at the lower site with 2 M added salt (figs. 2.26 and 2.27). Because the added electrolyte is decreasing the electrostatic contribution to binding, it is makes the non-ionic interactions, such as hydrogen bonding, relatively more important in determining the recognition of DNA by *N*-methylpyrrolecarboxamides, although it may also be altering the form of the DNA. Therefore, of the sites we have observed on the 167 bp and the 517 bp fragments, the one most efficiently recognized by P6E·Fe(II) is the 5′–ATTTTTAT–3′ site on the 517 bp fragment.

The broadening of the cleavage pattern could arise in three ways. The pyrrole compound could be sliding one or two base pairs in the groove. This effect is not seen under normal salt conditions and there is no reason to expect decreasing electrostatic interactions to produce it. The DNA may have changed conformation under the high salt conditions resulting in a larger set of bases accessible to the cleaving species. No sharp change in the cleavage patterns is

observed as would be expected for a DNA structure change and it is known that *N*-methylpyrrolecarboxamide compounds recognize B-form DNA and will convert A- or Z-form DNA to B-form upon binding.[17–22] The third possibility is that the negatively charged EDTA moiety is no longer held in a fixed position away from the DNA. With the higher salt to neutralize its negative charge, the EDTA does not have the strong repulsion from the phosphates causing it to be maximally extended from the DNA. The wider range of available delivery positions would make a greater number of bases accessible to the DNA cleaving species causing the cleavage pattern to be broadened.

DNA Binding Specificity. According to the n+1 rule and the bifurcated hydrogen bond model, the minimum recognition element for the *N*-methylpyrrolecarboxamide on B-DNA is two base pairs. If the recognition elements for the NH of the carboxamide are on adjacent residues on opposite strands of the helix, there are 10 bridged base possibilities: A-A, A-T, A-C, A-G, T-T, T-C, T-G, C-C, G-G, and C-G. The observed binding sites for the homologous series of poly-*N*-methylpyrrolecarboxamide compounds reveals that the preferences for bridged three center hydrogen bonds between adjacent bases on opposite helix strands decreases in the following order: A-T $\gg$ A-A $>$ T-T $>$ A-C, T-C, T-G $\gg$ A-G, C-G, C-C, G-G (table 2.4). From these data, the preferred two base-pair DNA recognition sequences for the *N*-methylpyrrolecarboxamide DNA binding unit are, in decreasing order, (5′–3′) TT $\gg$ TA, AT $>$ GA, GT, CT $\gg$ CA, CC, GC, CG.

Table 2.4: Bridged Hydrogen Bonds from Affinity Cleaving Data

Oligo-peptide	Bridged bases on opposite strands									
	517 bp Fragment									
DE	t - t	T - A	T - A	T - A	T - A	a - a	t - t			
P4E	t - t	T - A	T - A	T - A	T - A	A - A	t - t			
P5E	t - t	T - A	T - A	T - A	T - A	A - A	T - T			
P6E	T - T	T - A	T - A	T - A	T - A	A - A	T - T			
DE	t - t	a - a	A - T	T - T	A - A	A - T	t - t			
P4E	t - t	A - A	A - T	T - T	A - A	A - T	t - t			
P5E	t - t	A - A	A - T	T - T	A - A	A - T	T - T			
P6E	T - T	A - A	A - T	T - T	A - A	A - T	T - T			
	167 bp Fragment									
DE	t - c	t - a	a - a	A - T	A - T	T - T	T - A			
P4E	t - c	t - a	A - A	A - T	A - T	T - T	T - A			
P5E	t - c	T - A	A - A	A - T	A - T	T - T	T - A			
P6E	T - C	T - A	A - A	A - T	A - T	T - T	T - A			
	381 bp Fragment									
DE	a - a	g - t	a - a	A - T	A - T	T - T	T - A			
P4E	a - a	g - t	A - A	A - T	A - T	T - T	T - A			
P5E	a - a	G - T	A - A	A - T	A - T	T - T	T - A			
P6E	A - A	G - T	A - A	A - T	A - T	T - T	T - A			
	SV40 Fragment									
DE	T - A	T - A	T - A	T - A	t - a	t - a	t - a	t - t	a - a	t - a
P4E	T - A	T - A	T - A	T - A	T - A	t - a	t - a	t - a	a - a	t - a
P5E	T - A	T - A	T - A	T - A	T - A	T - A	t - a	t - t	a - a	t - a
P6E	T - A	T - A	T - A	T - A	T - A	T - A	T - A	t - t	a - a	t - a
P7E	T - A	T - A	T - A	T - A	T - A	T - A	T - A	T - T	a - a	t - a
P8E	T - A	T - A	T - A	T - A	T - A	T - A	T - A	T - T	A - A	t - a
P9E	T - A	T - A	T - A	T - A	T - A	T - A	T - A	T - T	A - A	T - A

In each sequence, bases contained in the site are symbolized with capital letters and the lowercase letters show the neighboring nucleotides.

The observation of G·C base pairs in the binding sites of these small molecules is something of a suprise. The crystal structure of netropsin bound

to a dodecamer of DNA has visualized the mode of A·T DNA recognition by *N*-methylpyrrolecarboxamides.[36-38] However, the fit of netropsin to the DNA is very tight in this crystal structure and it is not readily apparent how such an interaction could adjust to accommodate the guanine 2-amino group protruding into the minor groove. Perhaps the complex can adjust to remove the close van der Walls nonbonding contact between the pyrrole CH and the NH_2 group of guanine. As will be discussed subsequently, the recognition sequences containing G·C base pairs appear to be weaker binding sites and possibly the poly-*N*-methylpyrrolecarboxamide backbone simply is not fitting as deeply into the minor groove.

A more detailed statistical analysis of the binding site sequences is complicated by the fact that the number of unique sequences increases exponentially as the size of the binding site is increased. For a five base pair site, there are $4^5/2$ or 512 unique DNA sequences. For a six base pair binding site, there are $4^6/2$ or 2080 unique DNA sequences. In the 263 base pairs of DNA for which I have presented a high resolution analysis, 173 of the possible five base pair sequences are represented, about 34% of the total, but only 219 of the possible six base pair sites are observed, about 10% of the total. Because only DE·Fe(II) has representation of a significant fraction of its possible recognition sequences, it is the only affinity cleaving compound for which a more detailed analysis will be presented.

On the four DNA restriction fragments a total of 263 base pairs with 173 unique five base pair sequences have been studied and five strong and one weak

five base pair binding sites for DE·Fe(II) have been observed. The six DE·Fe(II) binding sites comprise only four unique recognition sequences (the 167 bp and 381 bp sites are the same as are one of the 517 bp and SV40 fragment sites) which is less than 3% of the unique 5 bp sequences assayed. This illustrates the high sequence specificity of the tri-*N*-methylpyrrolecarboxamide recognition element.

The highest affinity binding sites for DE·Fe(II) contain only A·T base pairs. For five base pair A·T sites there are $2^5/2$ or 16 unique sites. The four fragments studied here contain 15 of the 16 unique sites (table 2.5). DE·Fe(II) recognizes three of the A·T sequences strongly and a fourth one weakly. There is no readily apparent characteristic of this subset of recognized A·T sites which would explain their high affinity for tri-*N*-methylpyrrolecarboxamide compounds, although they have some interesting characteristics. One of the features of the strongest binding sites is that they tend to include few purine/pyrimidine switches, 5′–AAAAA–3′ on the 517 bp and SV40 fragments (0 switches) and 5′–AAATT–3′ on the 381 and 167 base pair fragments (1 switch). The preferred binding sites also tend to be included in longer runs of A·T base pairs. It may be that the stucture of a sequence within an A·T tract is better for polypyrrole binding. It has been shown that the groove width of G·C regions is wider than in A·T.[95] Perhaps the widening of the groove in a G·C region extends into the A·T region decreasing the hydrophobic contacts between the polypyrroles and the sugar residues of the DNA backbone.

Table 2.5: Occurance of 5 base pair A·T Sites on the 167, 381, 517, and SV40 Fragments

Site	167	381	517	SV40	Total
A A A A A	-	-	+	+	+
A A A A T	-	-	X	X	X
A A A T A	-	-	X	X	X
A A T A A	-	-	+	X	+
A T A A A	X	-	X	X	X
T A A A A	-	-	X	X	X
A T A T A	-	X	-	-	X
A T T A A	X	-	X	-	X
A T A A T	-	-	X	-	X
A A T T A	-	-	-	X	X
T T A A A	X	-	-	-	X
T A T A A	-	-	X	-	X
T A A A T	-	-	X	+	+
T A A T A	-	-	-	-	-
A A T A T	X	-	-	-	X
A A A T T	+	+	-	X	+
Total Sites	5/16	2/16	10/16	9/16	15/16

The X indicates the site is found on the fragment. The + indicates that the site is both found on the fragment and recognized by DE·Fe(II). The - indicates that the site is not found on the fragment.

These correlations can be used as a general rule of thumb for poly-*N*-methylpyrrolecarboxamide specificity, but it is important to remember that other factors influence this specificity as well. We have already seen that flanking sequences can influence the binding orientation of the small molecules at their binding site. In the SV40 restriction fragment, there is a run of eight consecutive adenine residues. Thus for DE·Fe(II) there are four possible binding sites containing five A·T base pairs with identical primary sequence, for P4E·Fe(II) there three identical six base pair poly(dA)·poly(dT) sites and for P5E·Fe(II)

there two identical five base pair poly(dA)·poly(dT) sequences; however, each of these compounds demonstrates well defined DNA cleavage patterns consistent with the recognition of a unique binding site. Therefore, all three of these compounds must be able to efficiently differentiate between two binding sites with the identical primary sequence. This differentiation must be based on the structural heterogeneity of the region produced by long-range interactions from outside the binding site sequence. Therefore, the DNA sequence specificity of the polypyrrole compounds is not dictated solely by the primary sequence of the binding site, but includes a combination of steric, electronic, and conformation effects.

Footprinting studies have detected other binding sites for the tri-*N*-methylpyrrolecarboxamide DNA recognition subunit which are not seen with the affinity cleaving compound. In addition, Crothers *et al.* reported a large cooperative interaction between distamycin and ethidium bromide binding to DNA.[34] To explore the possibility of allosteric binding with the footprinting reagent MPE·Fe(II) and poly-*N*-methylpyrrolecarboxamide compounds, affinity cleaving studies were carried out in the presence of a range of ethidium bromide concentrations. These studies showed no change in DNA recognition by DE·Fe(II). Another explanation for the difference in the number of binding sites observed with the two methods is a difference in binding site saturation between the two experiments. To examine the effect of binding site saturation with affinity cleaving reagents, reactions were carried out with increasing concentrations

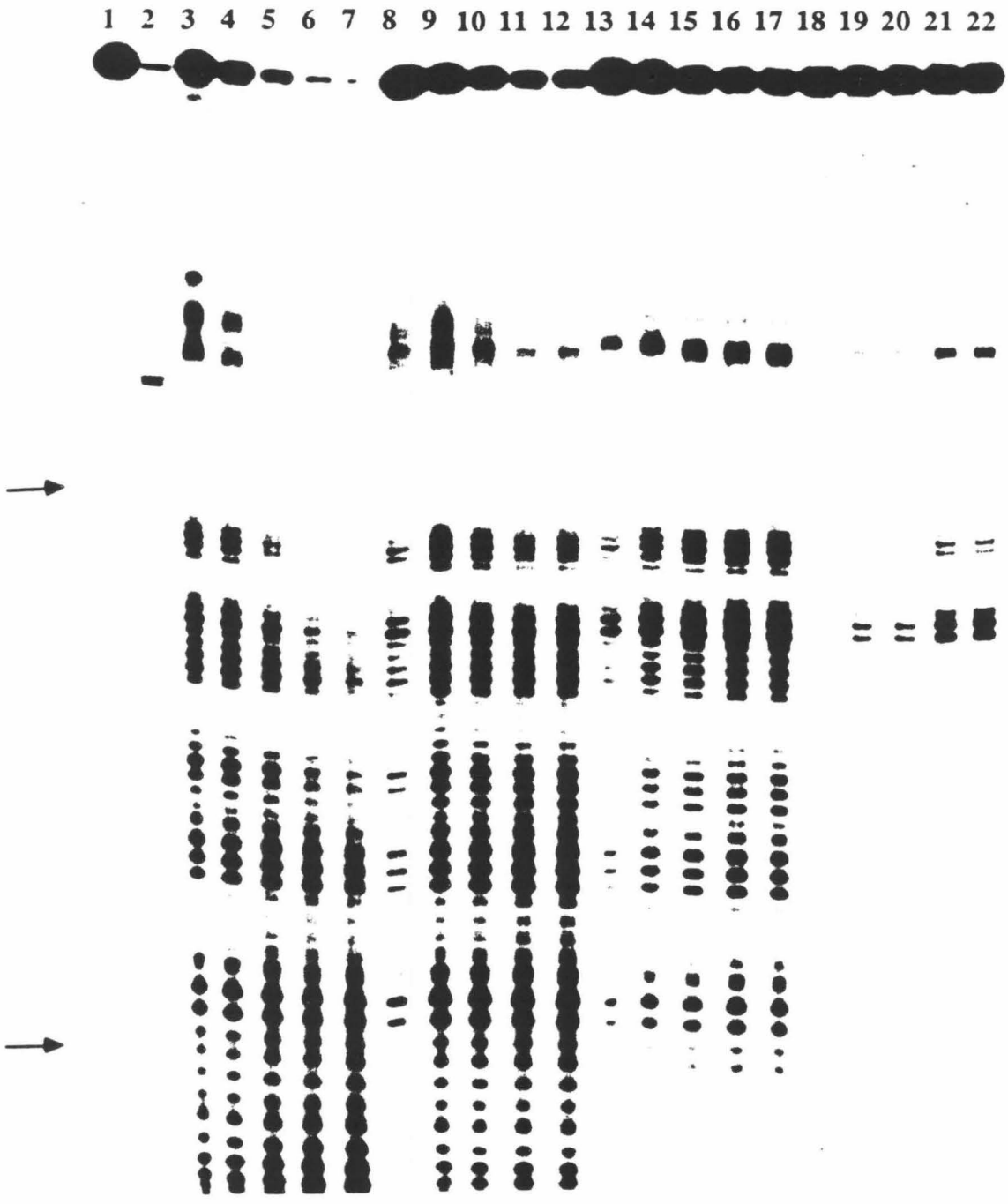
of P4E·Fe(II) using varied reaction times (fig. 2.28). These studies show that increasing concentrations of P4E·Fe(II) produce an increasing number of cleavage loci (fig. 2.29). This is consistent with a greater number of binding sites occupied at the higher polypyrrole concentrations establishing that it is the difference in binding site saturation which produced the apparent discrepancy in the observed binding sites between footprinting and affinity cleaving.

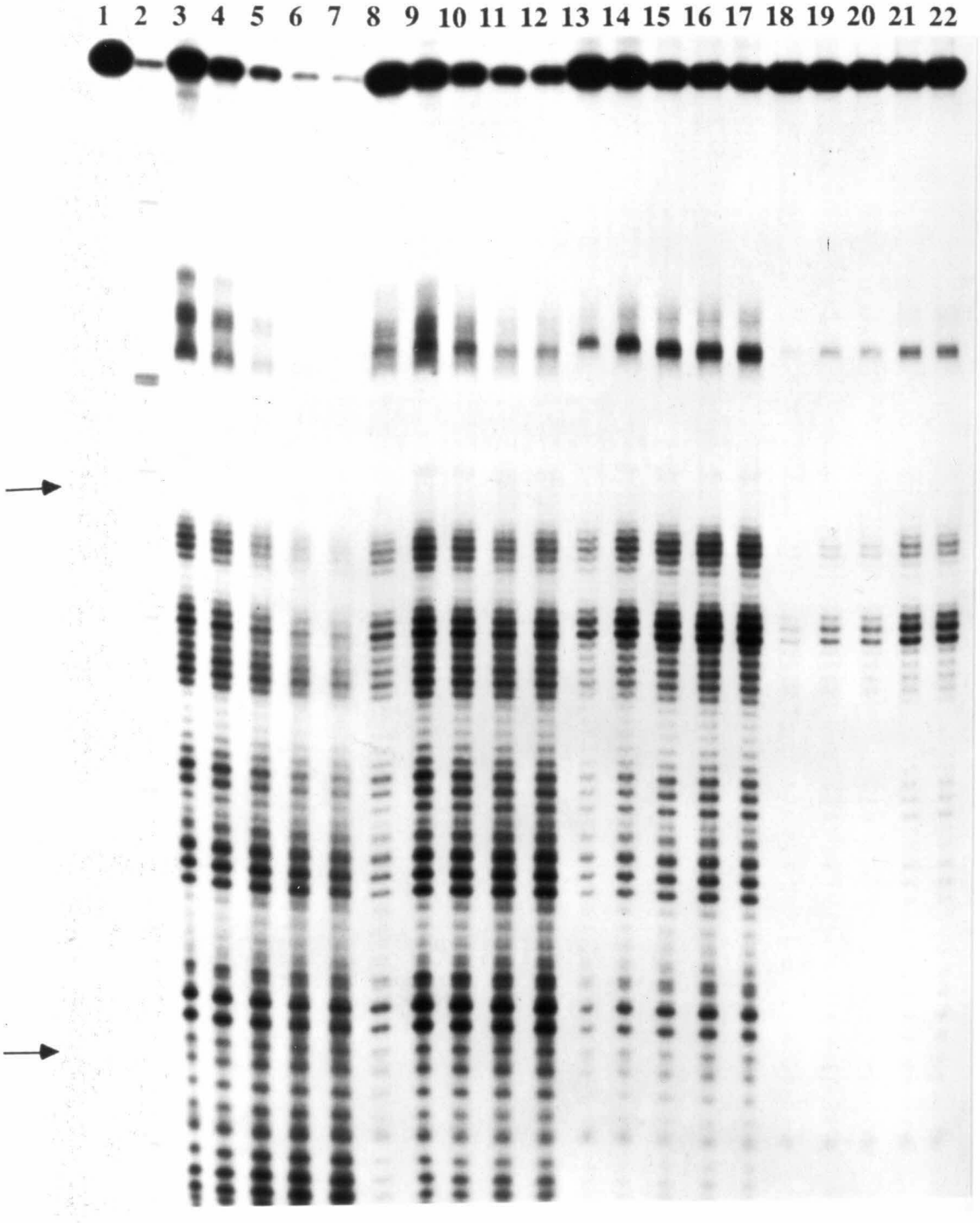
At the higher P4E·Fe(II) concentrations (40 μM, 5 min fig. 2.29), more sites are occupied, but all the sites show approximately the same amount of cleavage. This result could be interpreted as demonstrating that all sites have about the same affinity for the small molecule. Using a much lower P4E·Fe(II) concentration, it becomes apparent that one site, 5′–TAAATT–3′, has a much higher binding affinity than the others (5 μM, 1 hour fig. 2.29). Because the affinity cleaving method produces a positive signal on the autoradiogram, it can be used at much lower binding site saturation than footprinting allowing only the highest affinity sites to be detected.

These studies show that each of the tri-, tetra-, penta-, and hexa-, hepta-, octa-, and nona-*N*-methylpyrrolecarboxamide-EDTA compounds follow the $n+1$ rule binding one base pair more than their number of amides. This result is consistent with a common DNA recognition model in which each of the amides form bridging three-center hydrogen bonds between adjacent base pairs on opposite strands of the DNA. This result demonstrates that, up to at least nine *N*-methylpyrrolecarboxamide units, the DNA complex of these compounds does

Figure 2.28 Autoradiograms of high resolution denaturing gel electrophoresis of 3′ endlabeled 167 bp pBR322 restriction fragment (*Eco*R I-*Rsa* I) with varying P4E·Fe(II) concentrations and reaction times. The reactions were 100 μM in DNA base pairs (final concentration). The P4E·Fe(II) was allowed to equilibrate with the DNA for 1 hour at 37°C. Cleavage was initiated by the addition of DTT to a final concentration of 5 mM and allowed to proceed at 25°C for the time indicated. The reaction buffer was prepared to final concentrations of 40 mM Tris base and 5 mM sodium acetate and was adjusted to pH 7.9 with acetic acid.

Lane 1 is an intact DNA control lane after incubation under reaction conditions. Lane 2 is a chemical sequencing lane containing the products of a Maxam-Gilbert G-specific cleavage reaction. All reactions in lanes 3–22 were carried out with P4E·Fe(II). Lanes 3–7 contain 40 μM P4E·Fe(II); Lanes 8–12 contain 20 μM P4E·Fe(II); Lanes 13–17 contain 10 μM P4E·Fe(II); Lanes 18–22 contain 5 μM P4E·Fe(II). Lanes 3, 8, 13, and 18 were allowed to react 5 minutes at 25°C; Lanes 4, 9, 14, and 19 were allowed to react 15 minutes at 25°C; Lanes 5, 10, 15, and 20 were allowed to react 30 minutes at 25°C; Lanes 6, 11, 16, and 21 were allowed to react 1 hour at 25°C; Lanes 7, 12, 17, and 22 were allowed to react 1 hour at 25°C but without pre-eqilibration with the DNA.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

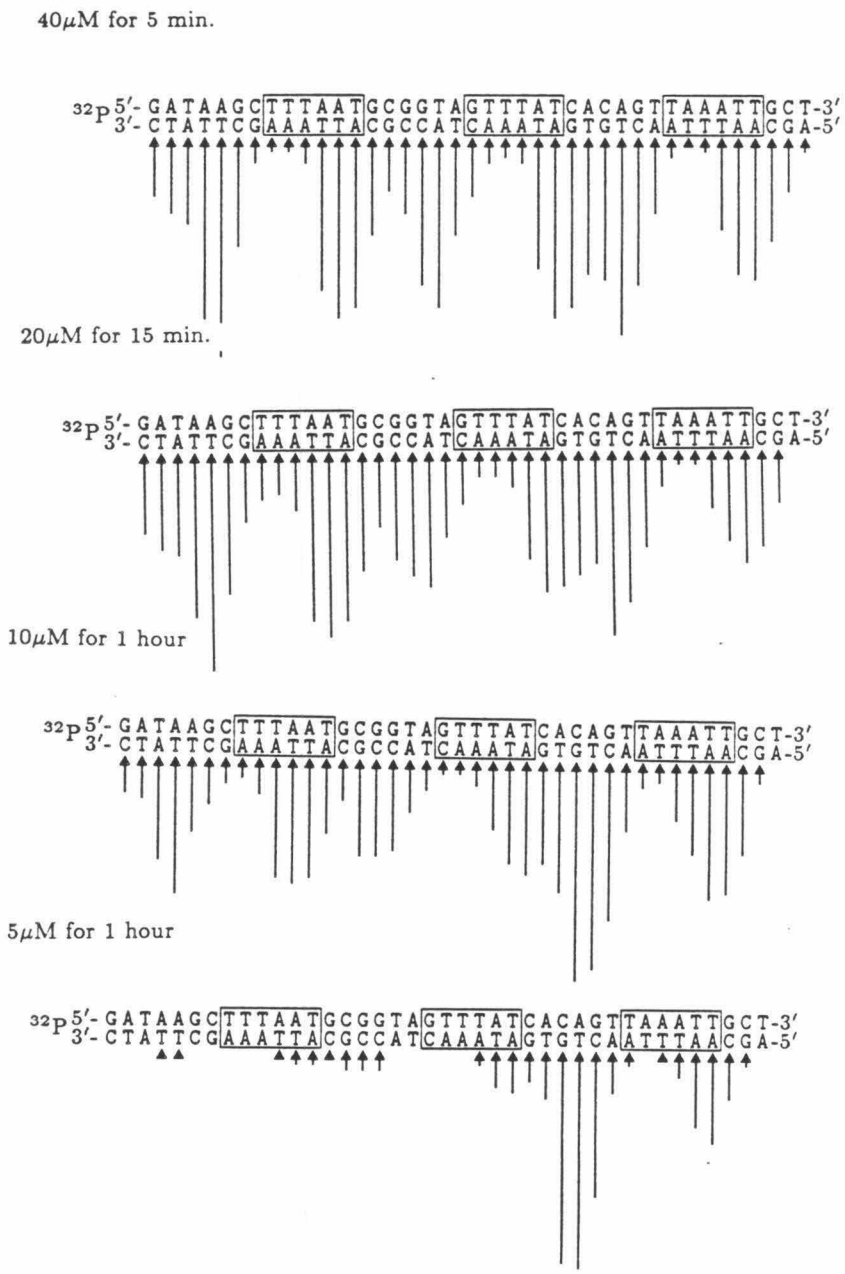


Figure 2.29 Histograms of the DNA cleavage patterns from cleavage of the 167 bp restriction fragment from shown in figure 2.28. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

not fall out of register with the DNA base pairs. The largest in this series of compounds, P9E·Fe(II), binds eleven base pairs, which is more than a complete turn of the B-form helix.

These studies also establish that the highest allinity sites for poly-*N*-methylpyrrolecarboxamide compounds on heterogeneous DNA are A·T sequences favoring those with a high polypyrimidine character. Their optimal binding sites are not solely dictated by the primary sequence of the binding site, but includes longer range influences from neighboring sequences. For compounds with binding site sizes larger than 5 bp, determining specificity by high resolution gels is impractical due to the exponential increase in unique sites. Further studies require a new assay that will permit the analysis of larger DNA fragments.

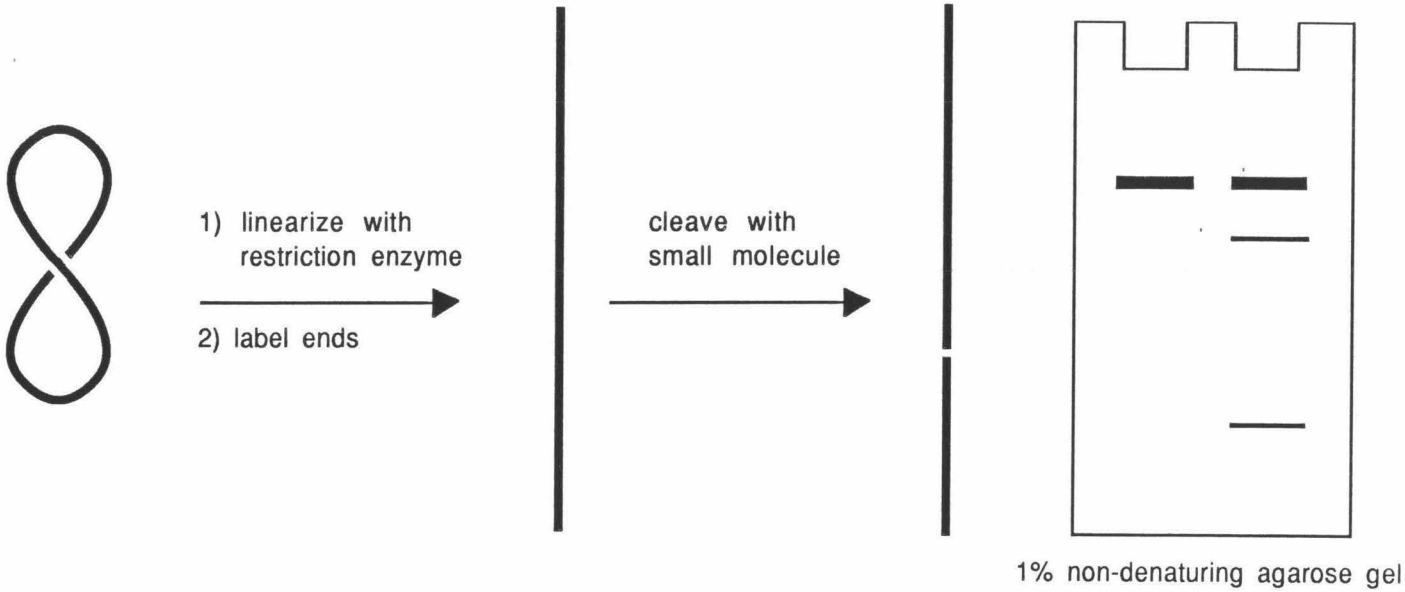
Large fragments of DNA were targeted because they allow simultaneous comparison of many potential binding sites. This facilitates the study of the relationship between the binding site size of a poly-*N*-methylpyrrolecarboxamide compound and the specificity with which it recognizes DNA. By assaying many potential binding sites, the optimum recognition sequences for the compound can be rapidly determined.

Large DNA is studied with a double strand cleavage assay requiring that both strands of the DNA be cleaved within a distance that allows denaturation of the intervening base pairs to produce two double strand DNA fragments (fig. 2.30). The reactions were run with linearized pBR322 in the same manner as the

reactions used for the high resolution gels. After the reactions were completed, they were electrophoresed on a non-denaturing agarose gel which separates double strand DNA based on size. If the double cleavage occurs randomly or with very low specificity, a continuum of fragments will be produced and no banding will be observed. On the other hand, with DNA cleaving agents of high specificity such as restriction enzymes a clear banding pattern will result. The number of bands produced by a compound is proportional to the frequency with which it finds its binding site; the greater the number of bands, the more often a recognition sequence occurs.

Studies with X-ray damaged DNA under a variety of salt and temperature conditions indicate that the two breaks must be within 3-24 bp.[96] Because this distance is so small, double strand breaks are assumed to result predominantly from two cleavage events at a single binding site and not two cleavage events at neighboring sites. However, this does not preclude the possibility that two or more distinct sites will be too close together to be resolved by the gel. These sites will then appear as one site with an average position on the DNA.

This type of experiment illustrates the power of the affinity cleaving method. The other commonly used method for determining the binding site of small molecules on heterogeneous DNA is footprinting which is capable of simultaneously assaying only about 100-200 base pairs. The affinity cleaving method can be used to simultaneously study thousands of base pairs and has been applied to DNA as large as 48 Kb (*vide infra*).



Scheme 2.5 The scheme for the double strand cleavage experiment. (Left) Supercoiled pBR322 is linearized with a restriction endonuclease. (Center) The DNA is labeled and reacted with an affinity cleaving compound under conditions which produce double strand cleavage. (Right) Non-denaturing agarose gel showing intact linearized plasmid and double strand cleaved plasmid.

The DNA used in these studies was pBR322 linearized by *Sty* I . This restriction endonuclease cleaves at an asymmetric recognition sequence on pBR322. This allows the DNA to be selectively radio-labeled at one or both ends with the appropriate choice of radio-labeled nucleotide using the standard 3′ labeling techniques outlined in the experimental section of this thesis. In these studies, CTP was chosen to label both ends of pBR322 and ATP was chosen to selectively label one end. The DNA was purified after labeling by preparative low melting agarose gel electrophoresis. Cleavage reactions were carried out under the same conditions as the high resolution gels described earlier and final concentrations, conditions, and reaction times are given in the gel figure legends.

The results of the double strand cleavage experiments are shown in figures 2.31 and 2.32. The most important initial observation is that all of the pyrrole compounds produce discrete banding patterns. Within this series of compounds, the three pyrrole compound, DE·Fe(II), shows the greatest number of bands on the gel and the total number of bands decreases as the number of pyrroles in the compound increases.

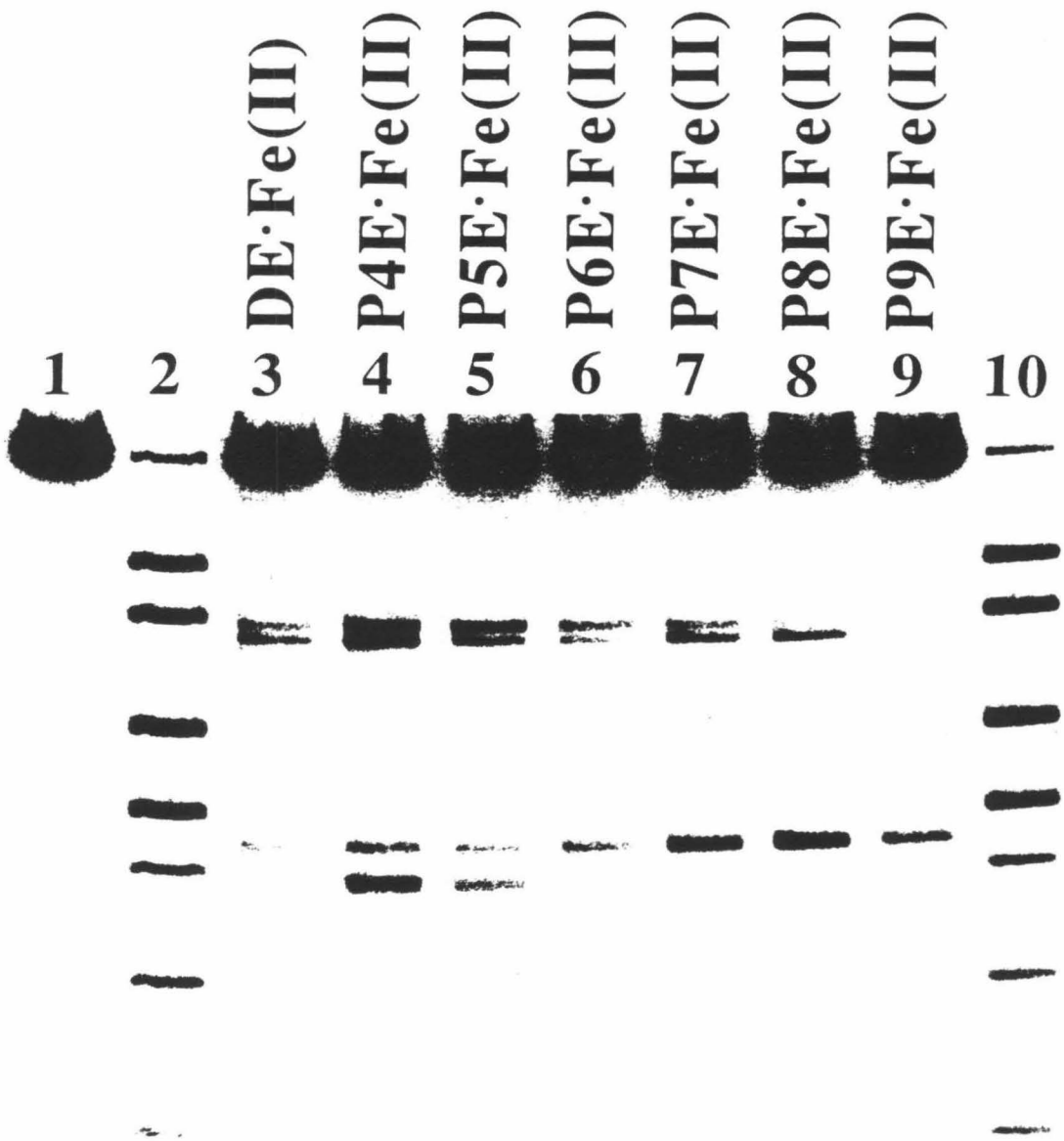
The molecular weights of the bands were determined by comparison with a molecular weight marker. For greater accuracy in molecular weight determination, some reactions were "spiked" with molecular weight standard. This provided a direct comparison of the mobilities of the reaction fragments with the molecular weight standard and avoided error introduced by lane to lane variations in electrophoretic mobility. Because each cleavage site produces two

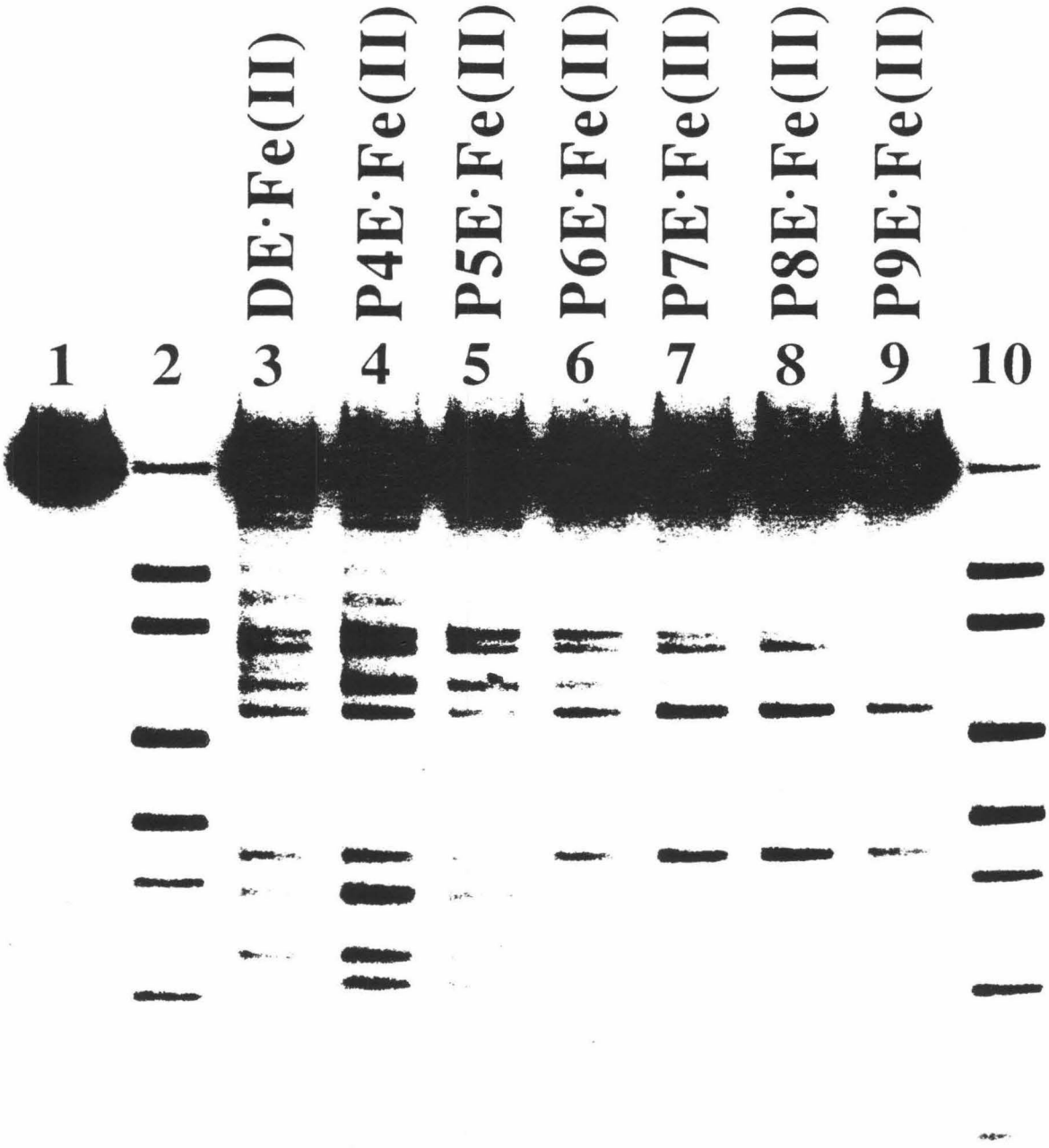
Autoradiograms of non-denaturing agarose gel electrophoresis of ^{32}P-end-labeled pBR322. The reactions were 100 μM in DNA base pairs (final concentration). The iron chelates of the compounds were allowed to equilibrate with the DNA for one hour at 65°C. Cleavage was initiated by the addition of DTT to a final concentration of 5 mM and allowed to proceed for 2 hours at 37°C. The reaction buffer was prepared to final concentrations of 40 mM Tris base and 5 mM sodium acetate and was adjusted to pH 7.9 with acetic acid.

Figure 2.31 pBR322 labeled at both ends of the *Sty* I site.

Figure 2.32 pBR322 labeled at one end of the *Sty* I sites

Lane 1 is an intact DNA control lane after incubation under reaction conditions. Lanes 2 and 10 contain molecular weight markers. Lane 3 is a DNA cleavage reaction with DE·Fe(II) at 1.2 μM; Lane 4 is a DNA cleavage reaction with P4E·Fe(II) at 0.8 μM; Lane 5 is a DNA cleavage reaction with P5E·Fe(II) at 0.6 μM; Lane 6 is a DNA cleavage reaction with P6E·Fe(II) at 0.6 μM. Lane 7 is a DNA cleavage reaction with P7E·Fe(II) at 0.6 μM; Lane 8 is a DNA cleavage reaction with P8E·Fe(II) at 0.6 μM; Lane 9 is a DNA cleavage reaction with P9E·Fe(II) at 0.6 μM.


DE·Fe(II)
P4E·Fe(II)
P5E·Fe(II)
P6E·Fe(II)
P7E·Fe(II)
P8E·Fe(II)
P9E·Fe(II)
1
2
3
4
5
6
7
8
9
10


DE·Fe(II)
P4E·Fe(II)
P5E·Fe(II)
P6E·Fe(II)
P7E·Fe(II)
P8E·Fe(II)
P9E·Fe(II)
1
2
3
4
5
6
7
8
9
10

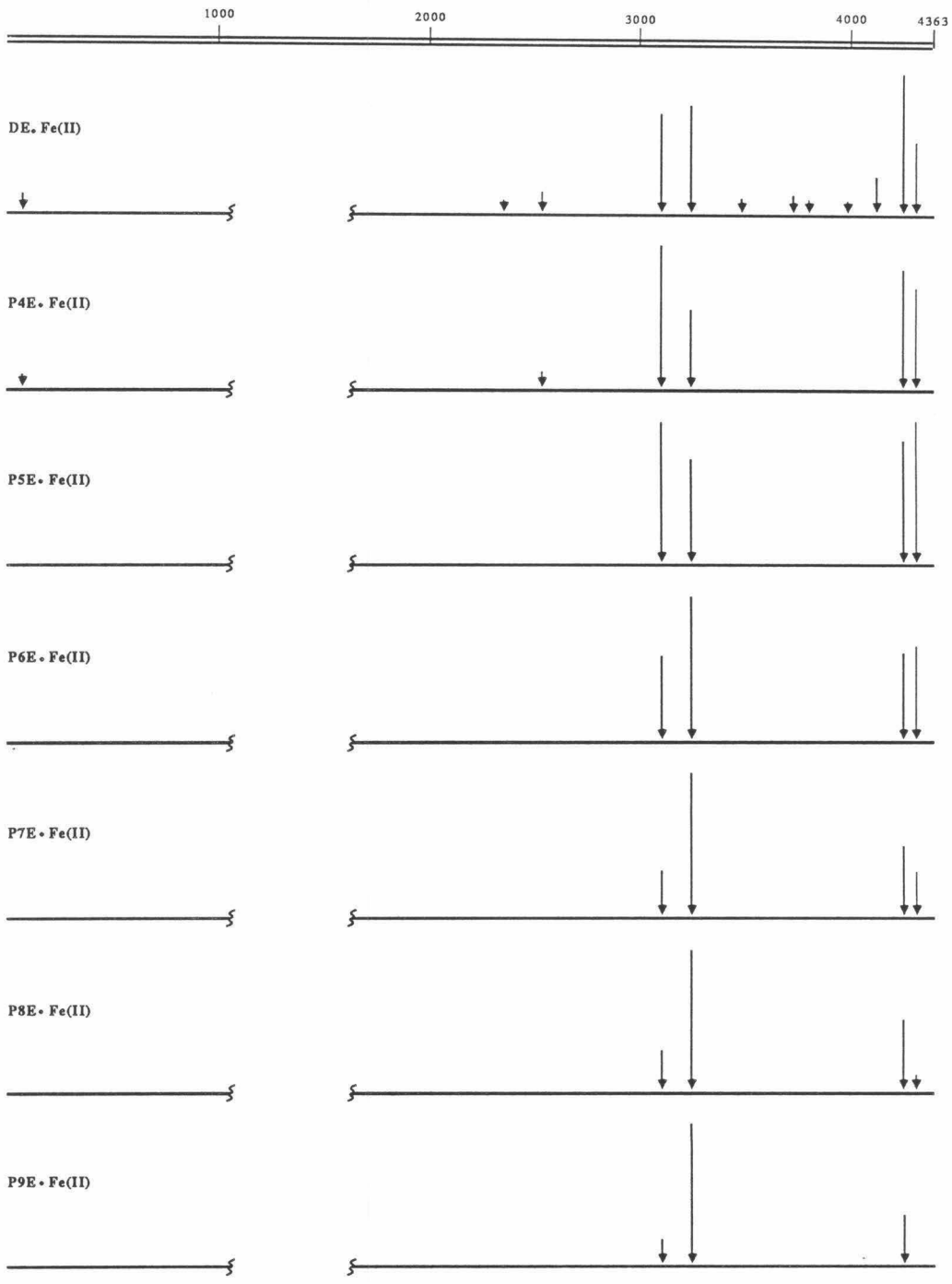


Figure 2.33 Histogram analysis of the gels in figures 2.31 and 2.32. Arrows represent the amount of cleavage observed at each location on pBR322. Broken line indicates the limit of resolution.

fragments of DNA which sum to 4367[97] the cleavage of double labeled pBR322 should produce symmetrical banding pattern where both fragments are labeled and are produced in equal amounts. This symmetry facilitates the pairing of the observed bands. In the series of reactions with singly labeled DNA, only one of the two fragments resulting from cleavage is radio-labeled. These data provide the information necessary to determine the direction to measure from the point of *Sty* I linearization towards the site of double strand cleavage.

The sites of cleavage were measured from the gel and compiled. Examination of the sequence of pBR322 at each of these sites found that with one exception there was a long run of A·T DNA within 5-30 bp with the strong sites more likely to be off by fewer bases (table 2.6). The exception is the site at 3855; however, this site is symmetrically positioned between several shorter runs of A·T DNA and may be a composite of unresolved bands resulting from cleavage at those sites. To examine the possibility that these A·T sites were the actual sites of DNA cleavage, the fragment lengths corresponding to cleavage of *Sty* I linearized pBR322 DNA at these sites were calculated. These values were then plotted against the band migrations measured from the gels and the correlation coefficient for the line was determined (r=0.9992). Comparison of this value with the correlation coefficient for the molecular weight marker lane (r=0.9991) in which all the molecular weights are known showed that these assignments were within the accuracy of the experiment. The results of this analysis were quantitated and converted into histogram form (fig. 2.33) and listed in tabular form (table 2.7).

Table 2.6: A Comparison of Measured and Proposed *N*-methylpyrrolecarboxamide Recognition Sequences on pBR322.

Measured Site	Proposed Site	Measured Site	Proposed Site
63	60	3695	3710
2335	2292	3855	3855
2531	2515	3975	3945
3068	3080	4130	4115
3215	3235	4260	4245
3460	3465	4343	4325

Specificity. For each position in a DNA nucleotide sequence there four possible bases which can recognize each other to form A·T and G·C complementary base pairs. With these constraints, the frequency of occurrence for a unique binding site of n base pairs is $(4^n)/2$ for odd n and $(4^n)/2+(4^{n/2})/2$ for even n (table 2.8). If the sequence contains only A·T base pairs, the possible permutations are given by $(2^n)/2$ for odd n and $(2^n)/2+(2^{n/2})/2$ for even n (table 2.8). The ratio of possible sites with n base pairs containing only A·T bases to those sites containing all four bases gives the frequency with which the A·T sites occur. From table 2.8, a five base pair A·T binding site such as that recognized by DE·Fe(II) occurs with a frequency of 3.1×10^{-2} or about once every 32 bp. It is remarkable then that DE·Fe(II) digestion of pBR322 produces a distinct banding pattern of only 12 distinguishable bands on the ~3700 bp analyzed. This result provides additional evidence that DE·Fe(II) does not simply recognize five consecutive A·T base pairs, but only a subset of these sequences.

The histograms state quantitatively what is qualitatively apparent from the gels themselves; increasing the number of *N*-methylpyrrolecarboxamides in

Table 2.7 Major Double Strand Cleavage Loci on pBR322

Site	Sequence	DE	P4E	P5E	P6E	P7E	P8E	P9E
60	CACAGTTAAATTGCTAACGC	W	W					
2335	CGTAAGGAGAAAATACCGCA	W						
2515	GAACCGTAAAAAGGCCGCGT	W	W					
3080	GGTGGTTTTTTTGTTTGCAA	S	S	S	S	M	M	W
3235	GATCCTTTTAAATTAAAAAT	S	S	S	S	S	S	S
3465	CAGATTTATCAGCAATAAAC	W						
3855	GTACTCAACCAAGTCATTCT	W						
3795	CAGCACTGCATAATTCTCTT	W						
3945	GCAGAACTTTAAAAGTGCTC	W						
4115	AATGCCGCAAAAAAGGGAAT	W						
4245	TGTATTTAGAAAAATAAACA	S	S	S	S	M	M	W
4325	TTAACCTATAAAAATAGGCG	M	S	S	M	M	W	

S, strong; M, moderate; W, weak.

Table 2.8:Binding Site Frequency for DNA Binding Molecules

Site Size (n)	Unique Sites (N)	Unique Site Frequency[a]	Unique A·T Sites	A·T Site Frequency[b]
1	2	5.0×10^{-1}	1	5.0×10^{-1}
2	10	1.0×10^{-1}	3	3.0×10^{-1}
3	32	3.1×10^{-2}	4	1.3×10^{-1}
4	136	7.4×10^{-3}	10	7.4×10^{-2}
5	512	2.0×10^{-3}	16	3.1×10^{-2}
6	2,080	4.8×10^{-4}	36	1.7×10^{-2}
7	8,192	1.2×10^{-4}	64	7.8×10^{-3}
8	32,896	3.0×10^{-5}	136	4.1×10^{-3}
9	131,072	7.6×10^{-6}	256	2.0×10^{-3}
10	524,800	1.9×10^{-6}	528	1.0×10^{-3}
11	2,097,152	4.8×10^{-7}	1,024	4.9×10^{-4}
12	8,390,656	1.2×10^{-7}	2,080	2.5×10^{-4}
13	33,554,432	3.0×10^{-8}	4,096	1.2×10^{-4}
14	134,225,920	7.5×10^{-9}	8,256	6.2×10^{-5}
15	536,870,912	1.9×10^{-9}	16,384	3.1×10^{-5}
16	2,147,516,416	4.7×10^{-10}	32,896	1.5×10^{-5}

a=1/N b=Unique A·T Sites/N

a molecule increases the specificity of the compound (fig. 2.33). Extending the pyrrole chain from 3 to 4 pyrroles decreases the number of cleavage loci observed from 12 to 6. The number of sites also diminishes moving from P4E·Fe(II) to P5E·Fe(II); however, further increases in the number of pyrroles in the molecules manifests an increase in DNA binding specificity primarily as a decrease in the cleavage intensity observed at the minor cleavage sites.

For P9E, one major cleavage site and two minor sites are observed. A unique 6 base pair binding site such as that recognized by a restriction enzyme occurs approximately every 2000 bp or with a frequency of about 5×10^{-4} (table 2.8). Therefore, on pBR322 a unique six base pair site should occur approximately twice. This is in qualitative agreement with the number of sites recognized by P9E, one strong and two weak, demonstrating that P9E, a synthetic organic molecule, shows at least formally the specificity of a restriction enzyme.

The sequences of the major cleavage sites are shown in table 2.7. There are six A·T stretches of 8 or more base pairs.[98] Of these six, only three are recognized as strong poly-*N*-methylpyrrolecarboxamide binding sites (table 2.9). The three sites not recognized by the pyrrole compounds contain a maximum of two consecutive pyrimidines while the three highest affinity sites have a minimum of 5 consecutive pyrimidines. This is consistent with earlier observations that the poly-*N*-methylpyrrolecarboxamide compounds preferentially bind recognition sequences of homopyrimidine A·T DNA.

A comparison of the cleavage at the 3235 and 4325 sites is shown in graphical form in figure 2.30. The site at 3235 contains the longest run of consecutive

Table 2.9: Sequences on pBR322 with ≥8 consecutive A·T base pairs.

Site	Sequence
	Recognized Sequences
3235	TTTTAAATTAAAAATGAAGTTTTAAAT*
4245	AAAAATAAA
4325	TATAAAAATA
	Sequences Not Recognized
3537	TATTAATT
4169	AATATTATT
4300	ATTATTAT

*This sequence contains 2 runs of ≥8bp which do not resolve on the agarose gel.[98]

A·T base pairs on pBR322 with 15. The site 4325 contains the second longest A·T sequence with 10 base pairs and few purine/pyrimidine switches. Both of these sites contain the sequence 5′–ATTTTTA–3′. The entire recognition sequence of DE·Fe(II) and P4E·Fe(II) as well as at least six of the seven base pairs in P5E·Fe(II)'s binding site have been shown to be included in this region in the high resolution gel studies. The 4325 site is found on the 517 bp fragment discussed earlier. Because the relative cleavage efficiency for DE·Fe(II), P4E·Fe(II), and P5E·Fe(II) are similar for both the 3235 and 4325 sites, they may be recognizing the same base sequence at both sites. As the binding site requirement is increased to eight base pairs with P6E·Fe(II), the sequence recognized at the two sites must be different and experimentally the efficiency of cleavage observed at these sites begins to diverge. As the pyrrole chain is further extended, the difference in recognition between the two sites continues to increase to a maximum at P9E·Fe(II).

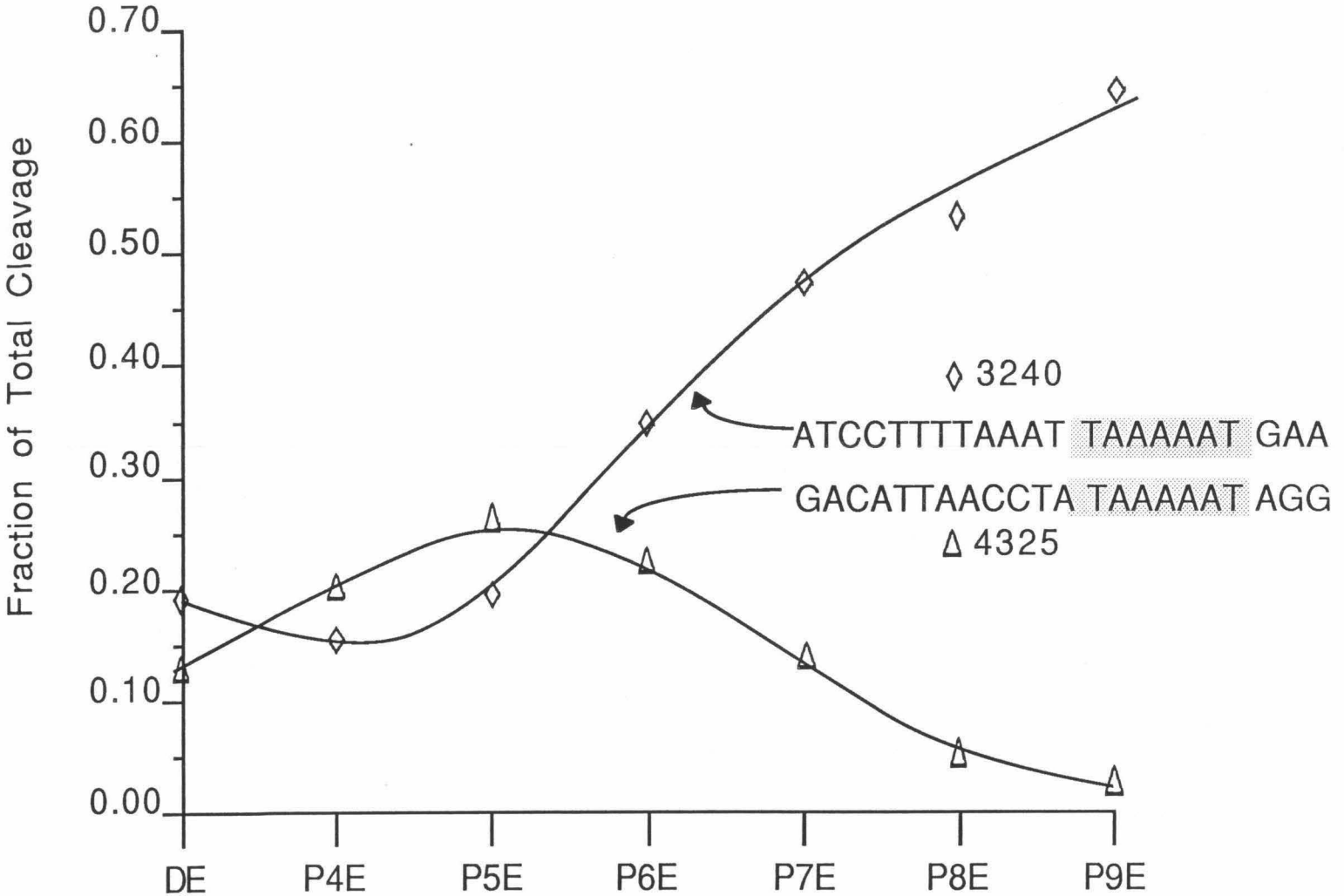


Figure 2.34 Comparison of the double strand cleavage of pBR322 at the 3235 and 4325 sites with DE·Fe(II)-P9E·Fe(II). Data points represent the fraction of the total cleavage observed for each compound at each site. The lines are provided to emphasize trends in the data points and are not intended to imply an analytical relationship between them.

Further examination of the sequence at the 4325 site shows that each additional base added to the recognition sequence beyond 5 bp is in an alternating purine/pyrimidine arrangment. At the 3235 site, the additional base pairs contain fewer purine/pyrimidine switches. The cleavage efficiency observed for these two sites again points to preferential recognition of homopyrimidine A·T regions by poly-*N*-methylpyrrolecarboxamides.

The cleavage efficiency of the polypyrrole-EDTA compounds at the 3080 site shows a maximum at P4E·Fe(II) and a steady decline in cleavage efficiency to P9E·Fe(II) (fig. 2.35). The sequence of the 3080 site contains seven consecutive adenines, the longest run of polypyrimidine A·T DNA on pBR322. It is the most efficiently recognized site on pBR322 for P4E·Fe(II) and is very efficiently recognized by P5E·Fe(II) (fig. 2.33). The reason that it is not the most efficiently recognized site for P5E·Fe(II), a compound containing 6 amides and binding 7 bp, may result from the flanking sequences exerting an influence which is not accommodating to *N*-methylpyrrolecarboxamide DNA recognition.

Time Course. Reaction time course studies with DE·Fe(II), P6E·Fe(II), and P9E·Fe(II) were carried out and the results of these studies are shown in graphical form in figure 2.36. The initial rate of cleavage is inversely proportional to the number of pyrroles in the molecule.

There are two mechanisms for double strand cleavage; in which the compound binds to the DNA and undergoes the redox cycle twice cleaving both strands and another where the compound binds the DNA, cleaves, *dissociates*,

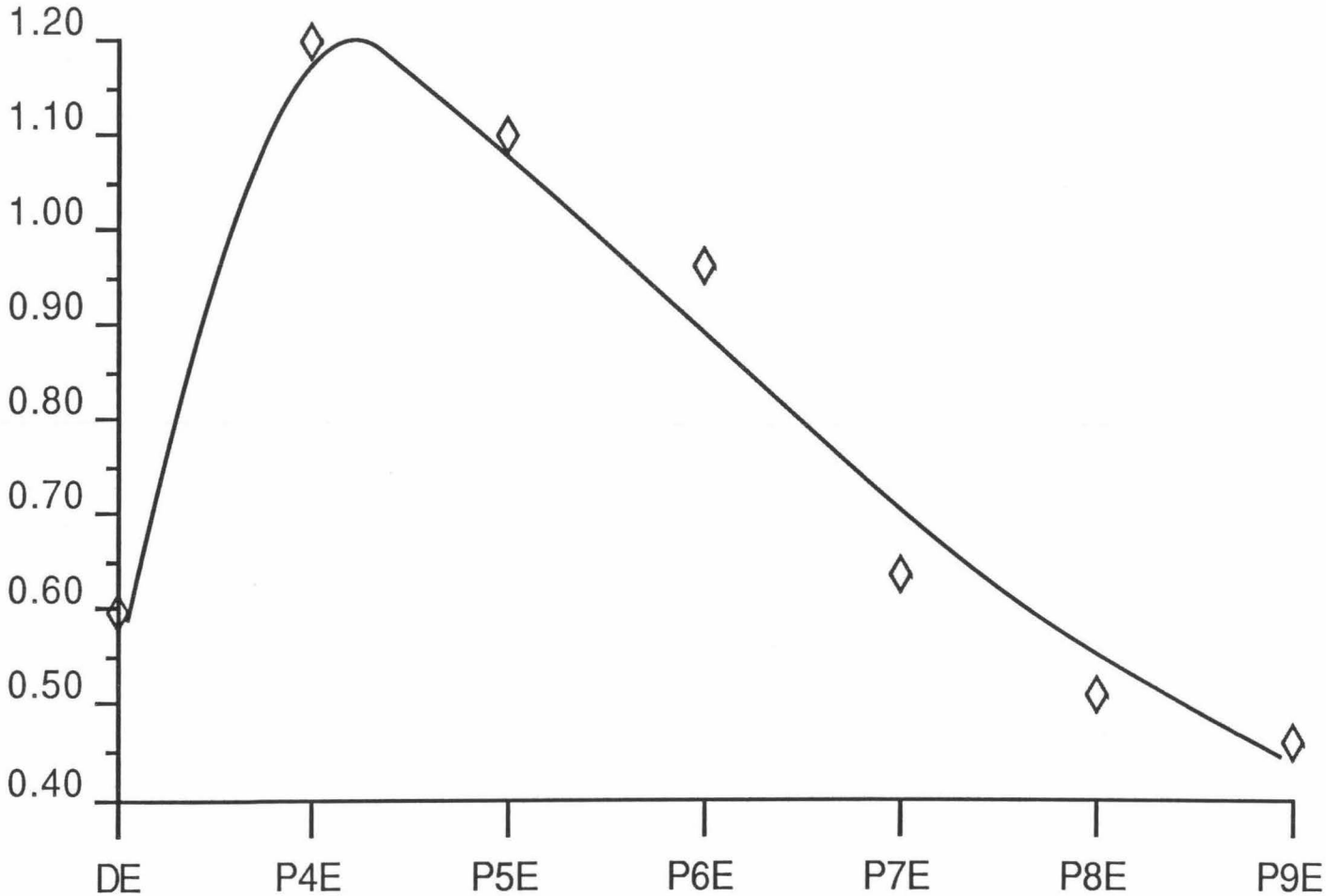


Figure 2.35 Comparison of the double strand cleavage of pBR322 at the 3080 site with DE·Fe(II)-P9E·Fe(II). Data points represent the fraction of the total cleavage observed for each compound at each site. The line is provided to emphasize trends in the data points and is not intended to imply an analytical relationship between them.

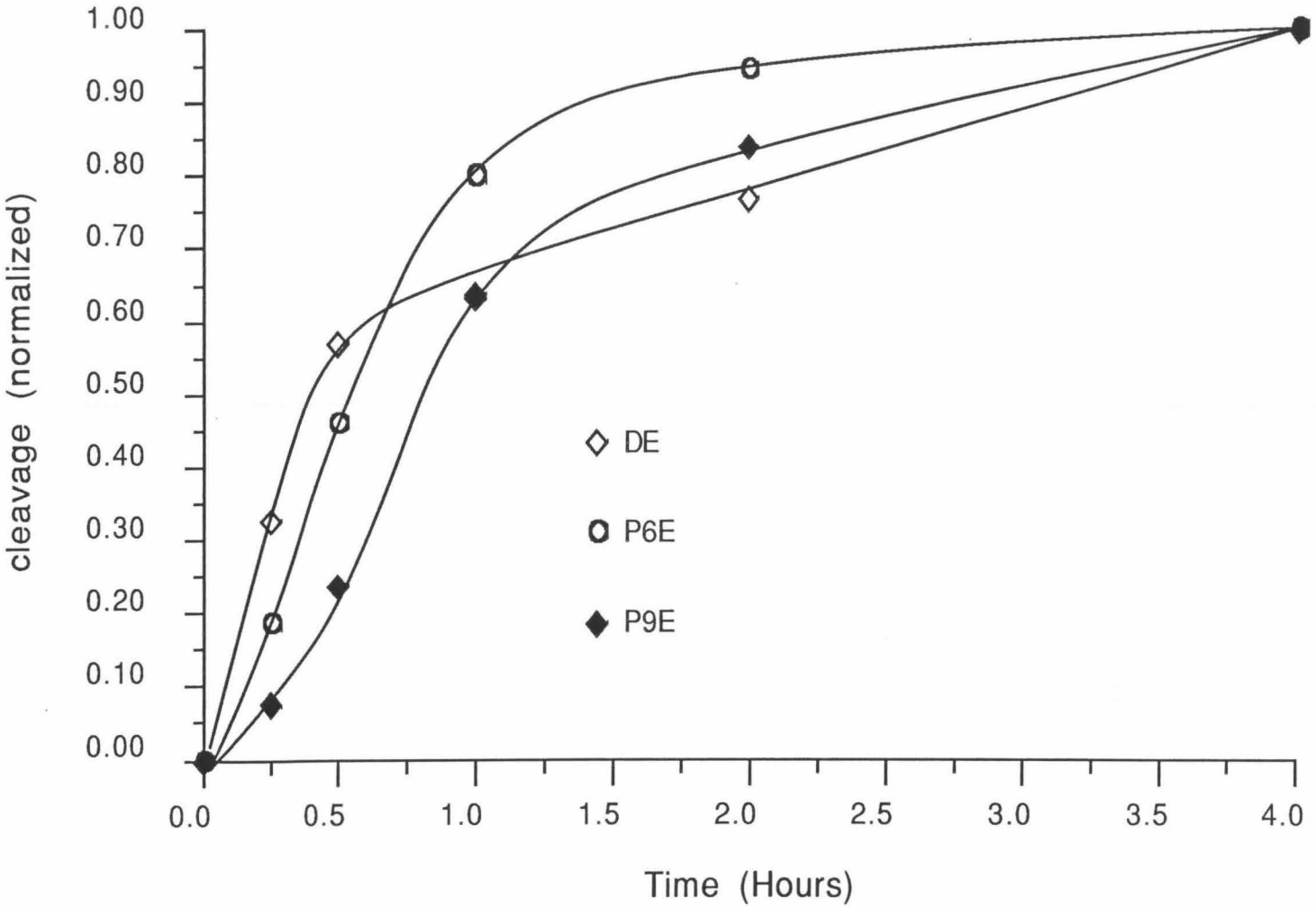


Figure 2.36 A time course study of the cleavage of pBR322 with DE·Fe(II)-P9E·Fe(II). Data points are ratios of the cleavage at the indicated time to the cleavage observed at four hours. The lines are provided to emphasize trends in the data points and are not intended to imply an analytical relationship between them.

is reduced, rebinds and cleaves the opposite strand. It is likely that the binding constant of the small molecules increases as the number of pyrroles in the molecule is increased or at least the kinetics of binding are slower with longer pyrrole chains.[61] Because the reaction path is so complicated, nothing quantitative can be said about the specific steps involved in double strand cleavage; however, the initial rates of double strand cleavage increase as the number of pyrroles in the molecule decreases consistent with the initial rate of cleavage being inversely related to the residence time of the molecules on the DNA. This result lends credence to the cleavage model which requires the polypyrrole compound to dissociate from the DNA and be reduced before returning to the binding site and cleaving the opposite strand. Inactivation of the poly-*N*-methylpyrrolecarboxamide-EDTA compounds may result from attack by the DNA cleaving species which could mediate pyrrole oxidation or EDTA degradation.

Pre-equilibration. For efficient DNA cleavage with the longer polypyrrole compounds, pre-equilibration with DNA is required before the initiation of cleavage. Pre-equilibration is typically carried out at 65°C for one hour. To address the importance of this step in DNA recognition, the following study with P9E·Fe(II) was carried out. P9E·Fe(II) cleavage of radio-labeled pBR322 DNA was run as usual except without pre-equilibration with the DNA. In another reaction, P9E·Fe(II) was equilibrated with unlabeled calf thymus DNA and after equilibration, radiolabeled restriction fragment and reducing agent were added.

In a third reaction, the DNA was equilibrated without P9E·Fe(II) and afterwards, P9E·Fe(II) and reducing agent were added. These result were compared to the control reaction with all components except reducing agent present in the reaction during equilibration and the results are tabulated in table 2.10.

Table 2.10: Effect of Pre-Equilibration with P9E·Fe(II)

$\frac{\text{Observed Cleavage}}{\text{Control Cleavage}}$	Conditions
13%	No equilibration
7%	P9E·Fe(II) added after equilibration
84%	^{32}P DNA added after equilibration

Without equilibration, DNA cleavage is inefficient and if the equilibration is carried out without P9E·Fe(II) the DNA cleavage is also low (table 2.10). Because both the reaction without equilibration and the reaction without P9E·Fe(II) show inefficient cleavage, the increase in cleavage efficiency observed upon equilibration is not solely dependent on a change in the DNA with heating. When the equilibration is carried out without the labeled DNA, cleavage is almost as efficient as when it is carried out with the intact reaction mixture. This demonstrates that both DNA and P9E·Fe(II) are needed to produce the increase in cleavage efficiency seen with equilibration. It also demonstrates that the effect is not due to an irreversible or very slowly reversible binding reaction between P9E·Fe(II) and the DNA because P9E·Fe(II) can exchange with the labeled DNA to produce cleavage which is almost as efficient as that observed with the intact reaction mixture. It is suprising that exchange is so facile because P9E·Fe(II) recognizes 11 bp of the DNA, about one B-form helical turn.

To exchange binding sites, P9E·Fe(II) must unwrap itself from one site and not only rewrap around another site, but locate an optimal site on DNA and then rebind.

The pre-equilibration effect could be explained by a simple aggregation of the compound in solution which is broken up by binding to DNA. UV studies with the pyrrole compounds follow Beer's Law demonstrating no indication of stacking of the aromatic pyrrole rings in solution. These studies were run over the concentration range of about 1–10 μM. The DNA cleavage reactions are typically run in the range of 0.1–1 μM.

Reorganization of the P9E·Fe(II) upon equilibration with DNA from an inactive conformation to one competent to bind DNA is also consistent with these results. A study was made of the entire series of compounds, DE·Fe(II)-P9E·Fe(II), comparing the efficiency of cleavage by each compound with and without equilibration. The results of this study are expessed as a ratio of cleavage observed without equilibration to cleavage observed with equilibration and are shown in graphical form in figure 2.33. An important feature of this graph is that there is an almost linear relationship between the increase in cleavage efficiency by equilibration and the number of pyrroles in the molecule with P9E·Fe(II) showing the greatest effect. This result is consistent with a polypyrrole reorganization dependent on the number of pyrroles in the molecule.

One conformational change which fits this criteria for the reorganization is the helical pitch of the molecule determined by the angle between adjacent

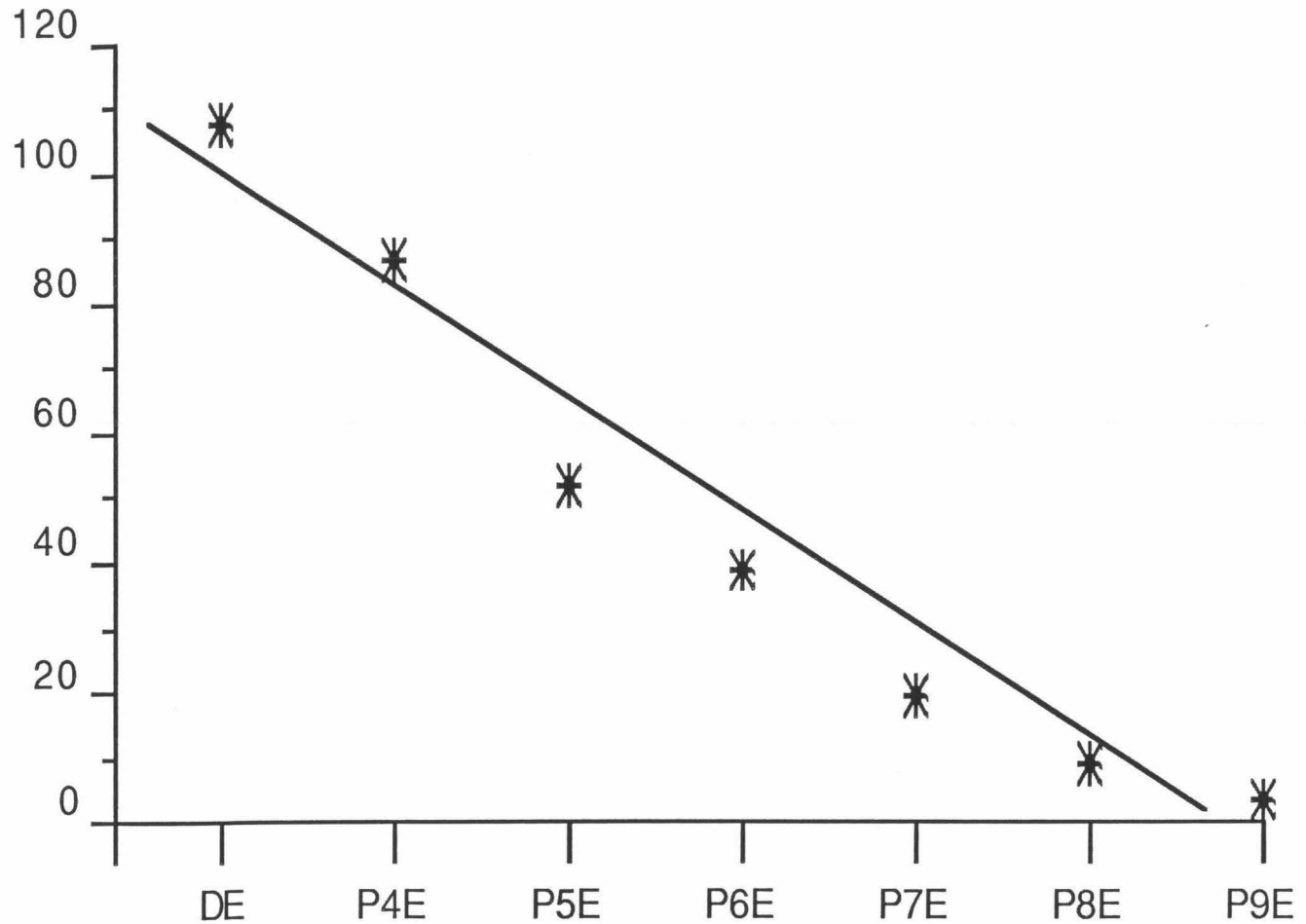


Figure 2.37 A study of the effect of pre-equilibration of the cleavage of pBR322 with DE·Fe(II)-P9E·Fe(II). Data points are ratios of cleavage without pre-equilibration to cleavage with pre-equilibration at 65°C for one hour. The line is provided to emphasize trends in the data points and is not intended to imply an analytical relationship between them.

pyrrole rings. It is known from the X-ray crystal studies of dipyrrole compounds that the pyrrole rings do not lie in a plane.[14, 15] This non-planarity gives the molecule a helical twist. B-form DNA is right-handed, so for recognition of DNA, the polypyrrole molecule must also adopt a right-handed conformation. If the angle between the pyrrole rings is in equilibrium, it is likely that the right or left handedness of the twist between any two pyrrole rings will be random. DNA could serve as a template to align the twist into a right-handed form. Alternatively, the DNA may act as a conformational sink binding the compound which has adopted the all right-handed conformation, driving the equilibrium in that direction. This mechanism requires that the conformational change is slow with respect to the rate at which the polypyrrole molecule exchanges between DNA molecules. This type of conformational change could explain the effect of equilibration on cleavage efficiency for the entire series of compounds because the ones with the most pyrroles would have the most angles to align and, therefore, would require the greatest reorganization. This model is also consistent with the pre-equilibration requirements observed with $(P4)_3E$ which are closer to those for P4E than what would be expected for compound containing 12-14 pyrroles and the observations with BED which behaves like DE.[71] This result establishes that it is not the total number of *N*-methylpyrrolecarboxamides in the molecule which dictates the pre-equilibration requirements but the number of contiguous pyrroles in each of the subunits suggesting a reorganization of the subunits as the limiting factor.

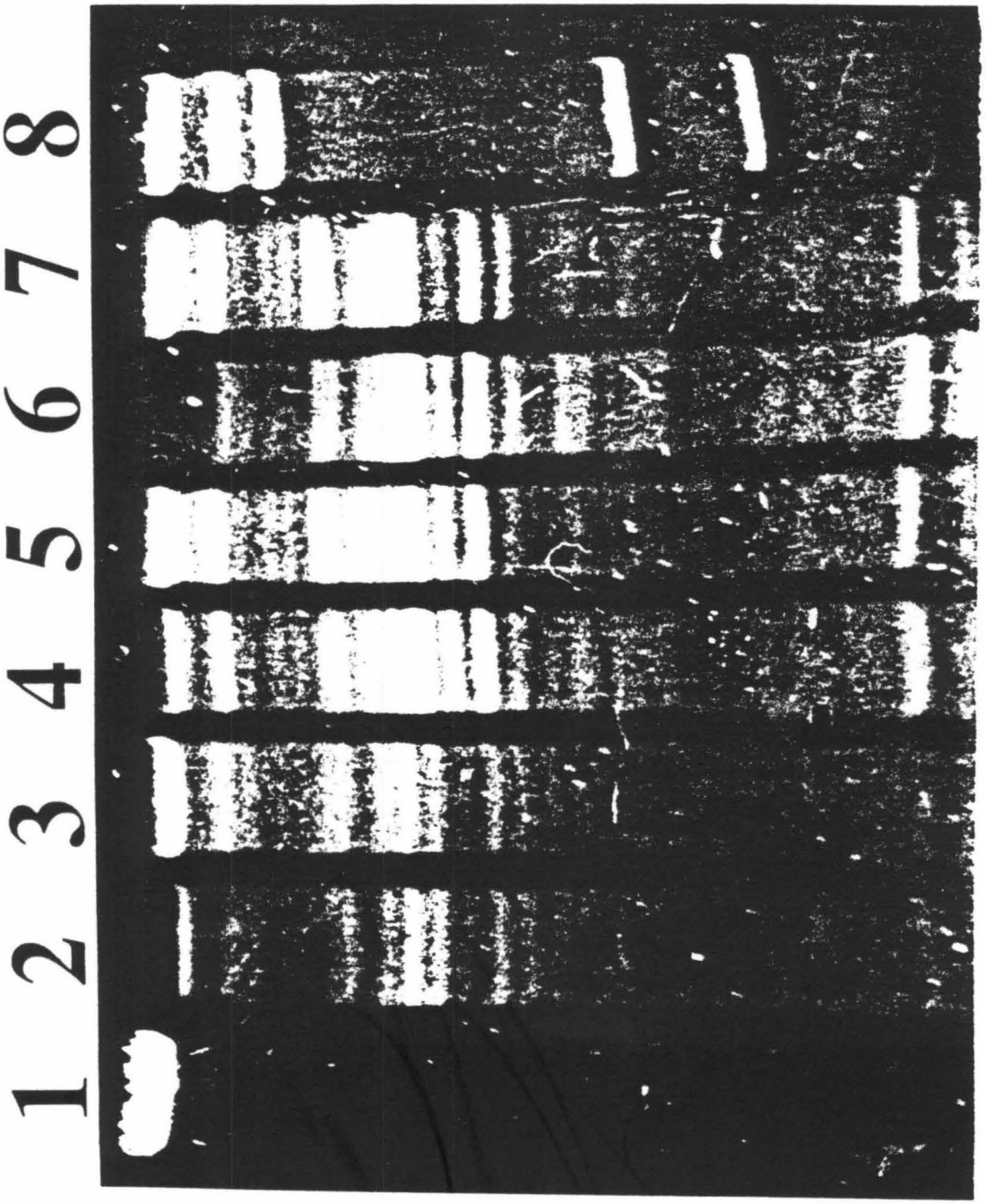
These double strand cleavage studies with pBR322 have established that the poly-*N*-methylpyrrolecarboxamide EDTA·Fe(II) compounds are capable of double strand cleavage of DNA in the kilobase range to afford a set of discrete DNA fragments. Preliminary studies with SV40 provide evidence that this property is general for DNA of this size range. In order to acertain whether the specificities of these compounds are great enough to allow them to selectively cleave larger DNA substrates, preliminary studies with λ-phage DNA were undertaken.

λ-phage is a 48,502 bp *E. Coli* bacteriophage. The reactions were run with 0.25 μg of DNA in 10 μL (~100 μM in DNA base pairs). Cleavage reactions were run as previously described for the high resolution gels. The reaction mixtures were loaded onto a 0.3% nondenaturing agarose gel and electrophoresed at 4°C with buffer recirculation. The DNA was stained with ethidium bromide, visualized by fluorescence and photographed. The results of this experiment are shown in figure 2.38.

This gel shows a distinct banding pattern with each of the three affinity cleaving reagents used in the study: DE·Fe(II), P8E·Fe(II), and BED·Fe(II). Increasing the size of the DNA recognition element increases the specific cleavage observed as evidenced by the sharpening of the bands in the P8E·Fe(II) lanes as compared with DE·Fe(II). It also increases the efficiency with which the DNA is cleaved as shown by the lower concentration of P8E·Fe(II) needed to digest the DNA. The third compound on this gel is BED·Fe(II), a dimer of two tri-*N*-methylpyrrolecarboxamide subunits which will be discussed in chapter 3. λ-phage DNA contains four stretches of 16 A·T base pairs as well as three longer

Figure 2.38 Photograph of a 0.3% non-denaturing agarose gel with λ-phage DNA stained with ethidium bromide and visualized by fluorescence. The reactions were 100 μM in total DNA base pairs (final concentration). The affinity cleaving compounds were allowed to equilibrate with the DNA for two hours at 65°C. Cleavage was initiated by the addition of DTT to a final concentration of 5 mM and allowed to proceed for 2 hours at 65°C.

Lane 1 Molecular weight marker lane containing the products of *Xho* I digestion and *Kpn* I digestion of λ-phage DNA as well as intact λ-phage DNA. Lane 2 contains λ-phage cleaved with BED·Fe(II) at 0.07 μM. Lane 3 contains λ-phage cleaved with BED·Fe(II) at 0.14 μM. Lane 4 contains λ-phage cleaved with P8E·Fe(II) at 0.28 μM. Lane 5 contains λ-phage cleaved with P8E·Fe(II) at 0.14 μM. Lane 6 contains λ-phage cleaved with DE·Fe(II) at 0.45 μM. Lane 7 contains λ-phage cleaved with DE·Fe(II) at 0.28 μM. Lane 8 is an intact DNA control lane incubated under reaction conditions.


1 2 3 4 5 6 7 8

stretches of 17, 19, and 23 base pairs. The observed cleavage is consistent with scission preferentially, although not exclusively, at these locations.

It is very suprising that DE·Fe(II) produces a banding pattern at all with λ-phage DNA. There are 512 unique 5 base pair sequences and even if we assume that DE·Fe(II) preferentially recognizes a unique 5 bp sequence it should find about 100 such sequences on λ-phage. Cleavage at these sites would result in 200 bands on the gel since each double strand cleavage event produces two DNA fragments. However, the distribution of these sites may not be random. If the resolution of the 0.3% agarose gel is ±500 bp, then if 10 high affinity sites occur within 500 base pairs the band produced will appear to be 10 times as strong as that produced from a 500 bp region containing only one high affinity site.

This points out another means for increasing the *apparent* specificity of the affinity cleaving compounds. Statistically, they will preferentially cleave the DNA in regions with a high concentration of binding sites causing their specificity to be greater than would be predicted from their recognition sequence alone. There is also evidence from biophysical studies for allosteric binding of poly-*N*-methylpyrrolecarboxamide compounds to DNA.[33, 34, 51, 99] Cooperative binding could enhance the binding affinity of the affinity cleaving compounds for regions with clusters of recognition sequences.

In practice this means that poly-*N*-methylpyrrolecarboxamide-EDTA·Fe(II) compounds may be capable of cleaving very large DNA on the order of chromosomes. One of the limiting factors in using the largest of the pyrrole compounds

with large DNA was that they had to be pre-equilibrated with the DNA at 65°C. These conditions are too rigorous for chromosomal DNA. The pre-equilibration studies with pBR322 suggest that it will be possible to equilibrate the affinity cleaving molecules with a smaller DNA and add that mixture to the large DNA. This would allow the entire reaction sequence to be carried out without ever raising the temperature of the chromosomal DNA above 37°C. The studies with λ-phage DNA suggest that these compounds may have the specificity required to cleave chromosomal DNA into discrete bands, preferentially cleaving in regions with a large number of recognition sequences. Cleavage of chromosomal DNA in A·T regions may be useful because enhancer and promoter regions which occur just before genes have been shown to be very A·T rich.

Summary. These studies have involved the synthesis and characterization of two series of poly-*N*-methylpyrrolecarboxamide molecules, one with EDTA attached and another with acetamide, where the number of contiguous *N*-methylpyrrolecarboxamide units has been incrementally increased. The effect of this incremental change on DNA recognition has been investigated at base pair resolution by both affinity cleaving and MPE·Fe(II) footprinting techniques. It was found that each incremental increase in the number of *N*-methylpyrrolecarboxamides in a molecule results in an increase the binding site size of one base pair. This observation led to a quantitative relationship between the number of amides in the molecule and the binding site size. This relationship is called the $n + 1$ rule and it states that a poly-*N*-methylpyrrolecarboxamide

molecule with n amides will bind $n + 1$ base pairs of DNA. This is consistent with the carboxamides of these compounds forming three center bridging hydrogen bonds between adjacent base pairs on opposite strands of the helix. The poly-*N*-methylpyrrolecarboxamide recognition element was found to preferentially bind poly dA·poly dT stretches; however, both binding site selection and orientation were found to be effected by flanking sequences. Using the affinity cleaving compounds on large pieces of DNA, it was demonstrated that increasing the binding site size increases the DNA binding specificity. Experiments with λ-phage DNA have shown that these compounds have the specificity required to efficiently cleave DNA into discrete bands at the 50 kb level and suggest that they may be able to operate on much larger DNA, possibly DNA of chromosomal size.

Chapter 3

Oligomers of Poly-*N*-methylpyrrolecarboxamide Compounds

The studies with the poly-*N*-methylpyrrolecarboxamide compounds have shown that increasing the number of contiguous pyrrole units in a molecule increases its DNA binding specificity. The number of *N*-methylpyrrolecarboxamides in a molecule may also be increased by oligomerizing polypyrrole subunits. The approach could be generalized to include the coupling of base-specific recognition elements of similar or diverse base pair specificities. With this strategy it could be possible to design molecules which are capable of recognizing DNA of any sequence.

Bisnetropsin derivatives linked in both a parallel and anti-parallel orientation have been reported.[62, 100] The tail-to-tail anti-parallel dimers employed hydrocarbon diacid linkers containing 4, 6, 8, or 12 carbons which were attached to the netropsin subunits through glycine residues. The interaction of the dimers with DNA was measured by circular dicroism (CD). From these studies, the authors concluded that both ends of the dimers interacted with the DNA. The binding site sizes observed by CD titration were: netropsin 5 bp; C_4 linker 7 bp; C_6 linker 10 bp; C_8 linker 10 bp, and C_{12} linker 11 bp. Comparison of saturation binding studies with poly(dA)·poly(dT) and calf thymus DNA indicated that the dimers bound the DNA more specifically that the monomer, netropsin. The dimers were also found to bind the DNA more strongly through competitive binding between the C_8 dimer and a flourescent netropsin analog. This study concluded that the binding constant of the C_8 dimer was at least 2 orders of magnitude greater than that of the netropsin monomer (5×10^6).

Schultz and Dervan prepared an affinity cleaving bisdistamycin compound.[101] This molecule, bis(EDTA-distamycin) (BED), was a tail-to-tail dimer of ED·Fe(II) which utilized a flexible linking element, heptanedioic acid. BED·Fe(II) was found to efficiently cleave supercoiled pBR322 at nanomolar concentration, almost three orders of magnitude lower than the monomer ED·Fe(II). High resolution gel studies revealed both monomeric and dimeric binding sites. This demonstrated that for BED·Fe(II) monomeric binding was competitive with dimeric binding. Recent footprinting studies with the bisnetropsin dimers indicate that these molecules also bind DNA in monomeric as well as dimeric binding modes.[102]

BEDF·Fe(II). In an effort to overcome the problem of competing monomeric and dimeric binding modes with bis-*N*-methylpyrrolecarboxamide compounds, the linking element was redesigned. Important considerations for both the linking and recognition elements are that they must be compatible with the same groove (e.g., minor or major) and conformational state (e.g., B, A, Z) of the DNA. Fumaramide was chosen to link two ED subunits because it is more rigid than heptanedioic acid and can more closely mimic the *N*-methylpyrrolecarboxamide in size and shape (fig. 3.1). This dimerization created the sequence-specific DNA binding molecule bis(EDTA-distamycin)fumaramide **31** (BEDF) (fig. 3.2).[103]

Synthesis. The tripyrrole nitro amine **32** is available from the tripyrrole nitro acid **21** using a procedure reported by Schultz (scheme 3.1).[71, 101] The primary

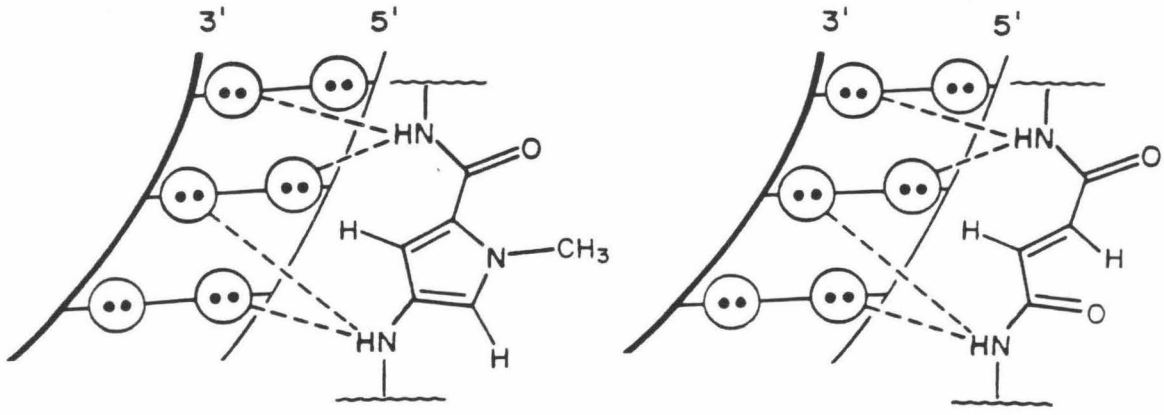


Figure 3.1 Circles with two dots represent lone pairs of electrons on N3 of adenine and O2 of thymine at the edges of the base pairs on the floor of the minor groove of the right-handed B-DNA helix. Dotted lines are bridged hydrogen bonds to the amide NHs. (Left) Model of *N*-methylpyrrolecarboxamide binding two base pairs in the minor groove at A·T DNA. (Right) Model of fumaramide binding two base pairs in the minor groove of A·T DNA.

amine on **32** was protected as the *tert*-butyl carbamate with di-*tert*-butyl dicarbonate. The terminal nitro group of the *t*-boc tripyrrole compound **33** was hydrogenated under atmospheric pressure and the resulting amine was dimerized on fumaric acid using dicyclohexylcarbodiimide and *N*-hydroxybenzotriazole as the coupling reagents to afford the *t*-boc protected dimer **34**. Deprotection of the primary amines with anhydrous TFA/CH_2Cl_2 (1:2) followed by acylation with the imidazolide of EDTA triethyl ester afforded BEDF-hexaethyl ester **35**. The ethyl ester groups were removed by hydrolysis with lithium hydroxide to yield the final compound, BEDF **31**.

DNA recognition by BEDF·Fe(II) was examined on two ^{32}P-endlabeled restriction fragments from plasmid pBR322: the 167 bp fragment, *Eco*R I-*Rsa* I

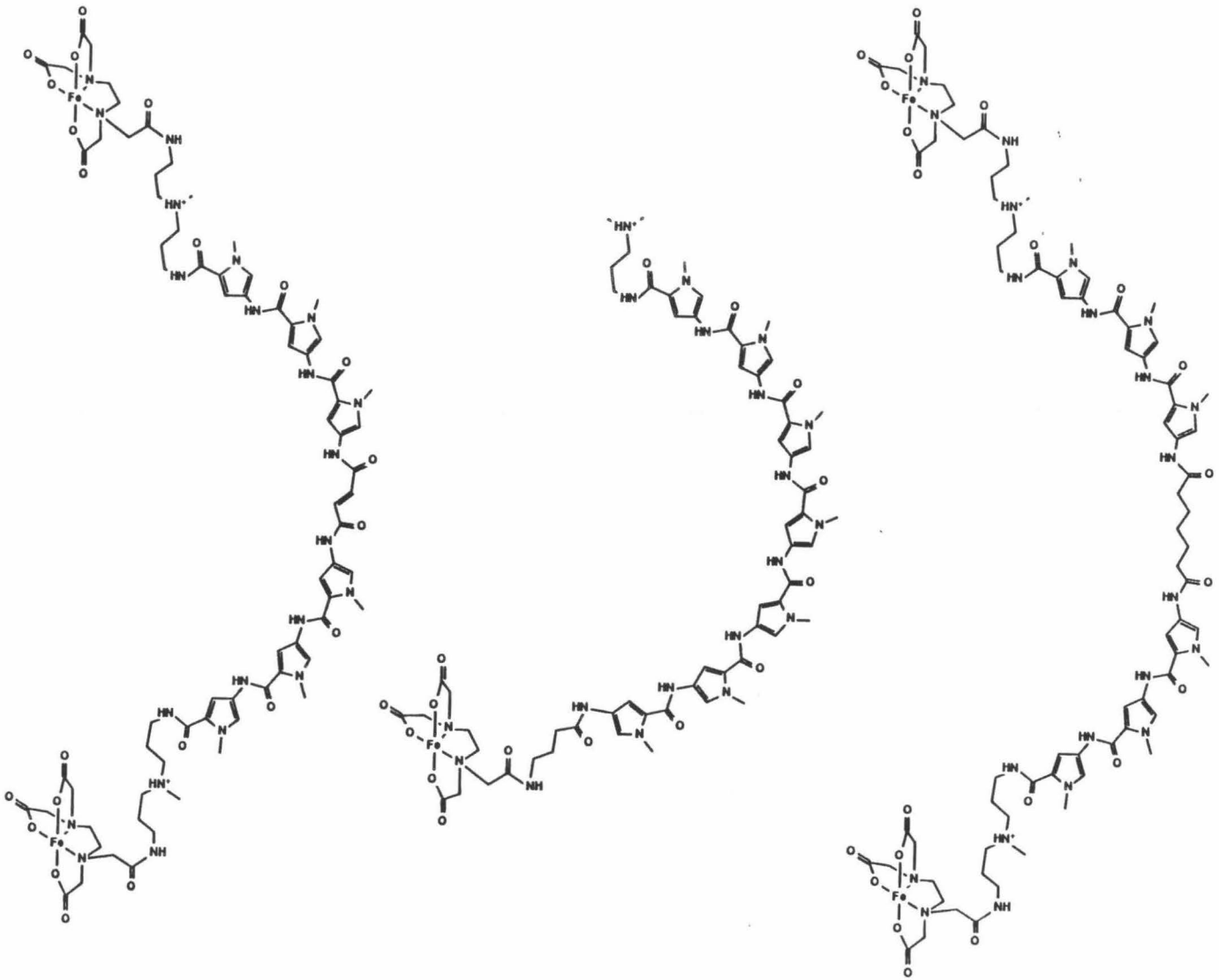
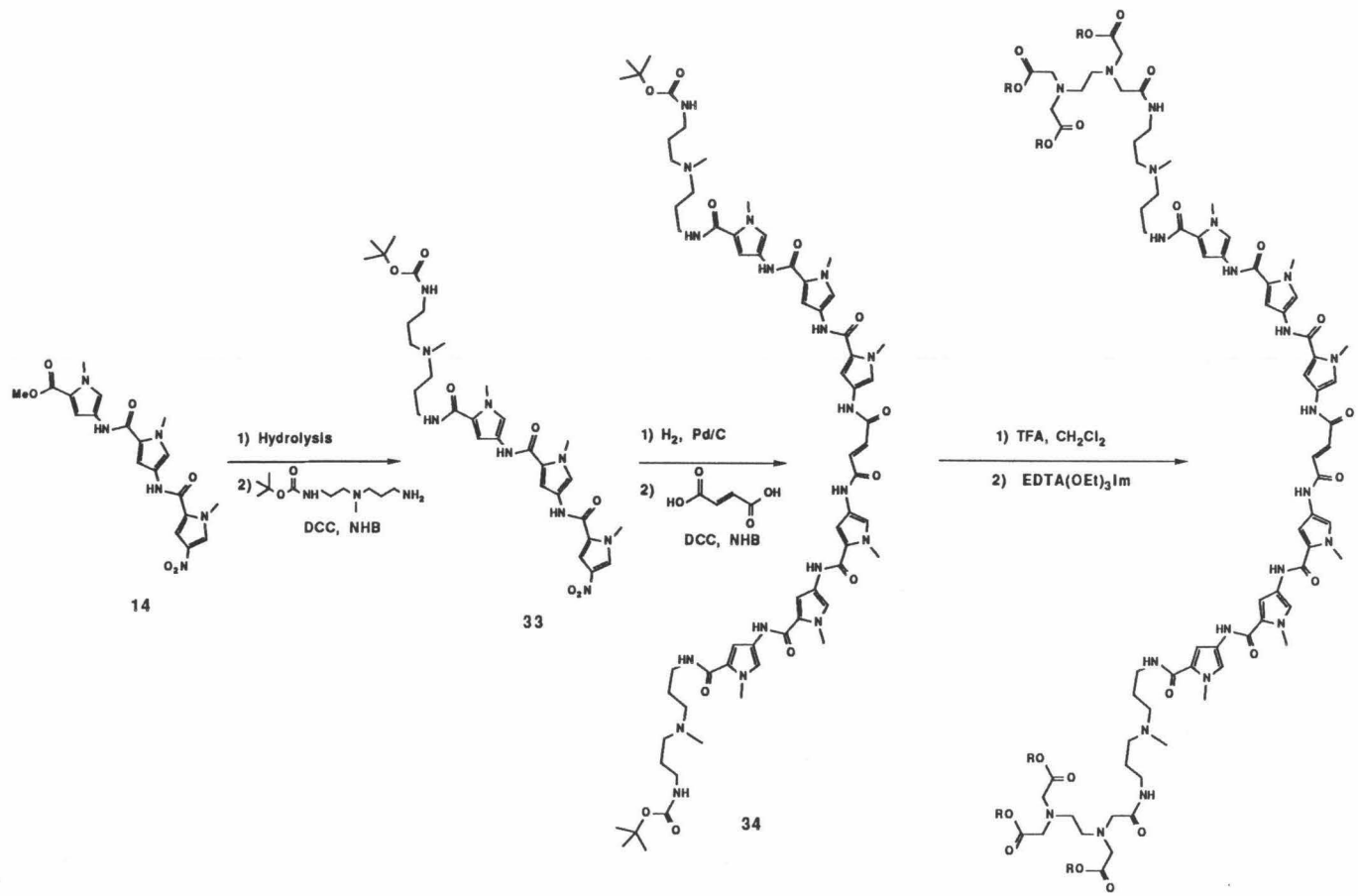


Figure 3.2 (Left) BEDF·Fe. (Center) P7E·Fe. (Right) BED·Fe.



Scheme 3.1 Synthesis of BEDF

(1-167); and the 517 bp fragment, *Eco*R I-*Rsa* I (4363-3847). For comparison the hepta-*N*-methylpyrrolecarboxamide-EDTA compound P7E·Fe(II) was also studied. Subsequent work optimized equilibration and reaction conditions for BEDF·Fe(II) making cleavage more efficient. Cleavage of the 517 bp fragment by BEDF·Fe(II) under the optimized condition is also presented. The cleavage reactions were carried out as described earlier and the final reaction conditions and concentrations are given in the figure legends.

167 bp Fragment. This fragment shows a complex cleavage pattern for BEDF·Fe(II) at the sequence 5′–GTTAAATTGCTA–3′ (base pairs 55–66 on pBR322). Analysis of the pattern is consistent with two overlapping binding sites (fig. 3.8). These two sites are assigned sizes of 8 and 9 base pairs, respectively. For P7E·Fe(II), one clearly resolved 9 bp binding site at the sequence 5′–AAGCTTTAA–3′ (base pairs 29–37 on pBR322) is observed. A second binding site for P7E·Fe(II) is also seen in the same region as the overlapping binding sites for BEDF·Fe(II). This cleavage is interpreted as a single 9 bp binding site.

517 bp Fragment. Two binding sites were observed for BEDF·Fe(II) on this fragment, 5′–ATTTTTATA–3′ and 5′–ATAATAAT–3′ (base pairs 4330–4322 and 4308-4301 on pBR322, respectively). The site at base pairs 4330-4322 is the stronger of the two and is assigned a binding site size of 9 base pairs. The weaker site at base pairs 4308-4301 is assigned a binding site size of 8 bp. P7E·Fe(II) has only one binding site in this region and it is assigned a binding site size of 9 base pairs (base pairs 4330-4322 on pBR322).

Figure 3.3 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 167 bp pBR322 restriction fragment (*Eco*R I-*Rsa* I). The reactions were 100 μM in total DNA base pairs (final concentration). The iron chelates of the compounds were allowed to equilibrate with the DNA for one hour at 25°C. Cleavage was initiated by the addition of DTT, to a final concentration of 5 mM, and allowed to proceed one hour at 25°C.

Odd numbered lanes are 5′-endlabeled DNA. Even numbered lanes are 3′-endlabeled DNA. Lanes 1 and 2 are intact DNA control lanes after incubation under reaction conditions. Lanes 3 and 4 are chemical sequencing lanes containing the products of Maxam-Gilbert G reactions. Lanes 5 and 6 are DNA cleavage reactions with BEDF·Fe(II) at 2.5 μM; Lanes 7 and 8 are DNA cleavage reactions with P7E·Fe(II) at 6.6 μM.


1 2 3 4 5 6 7 8

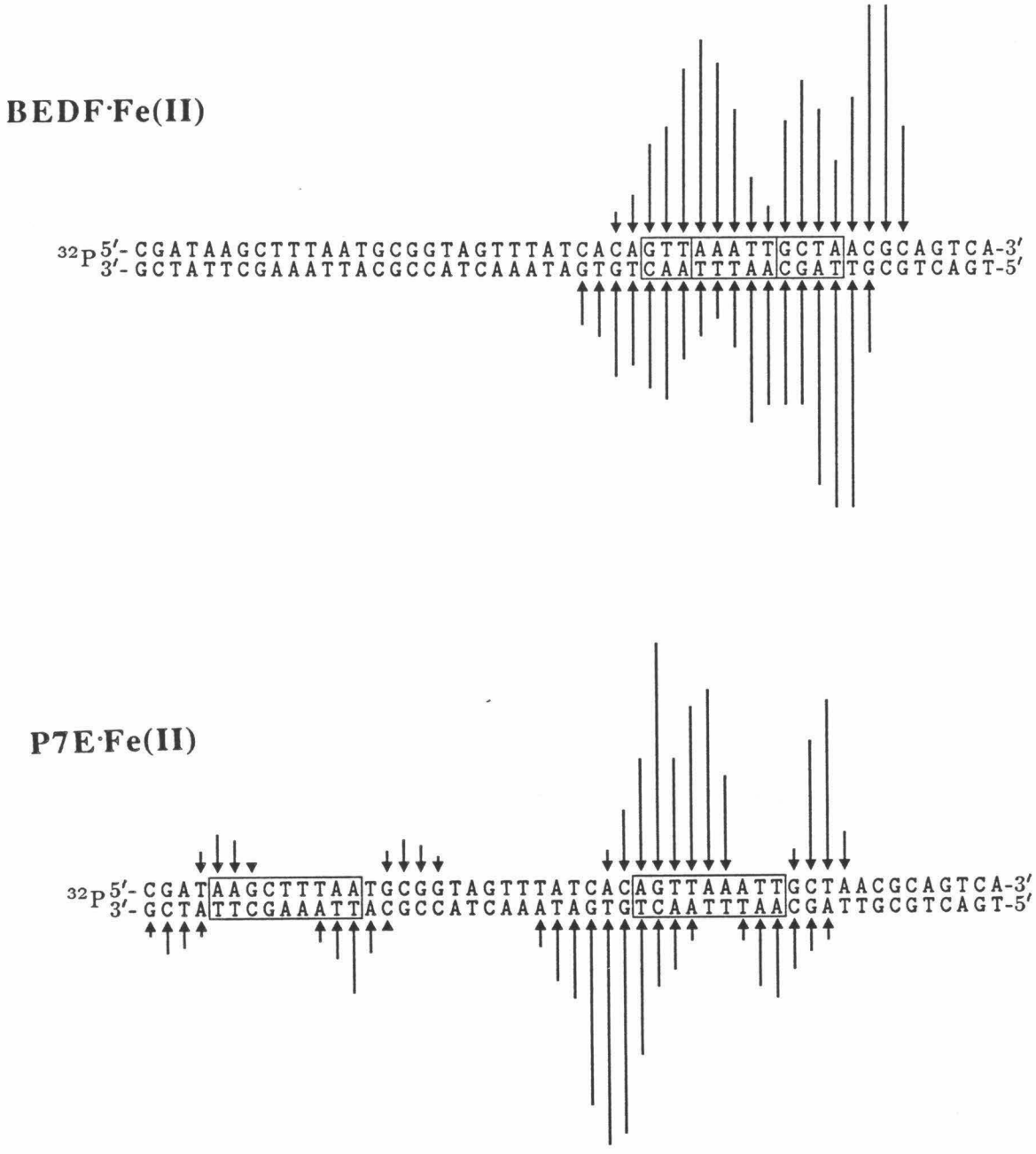
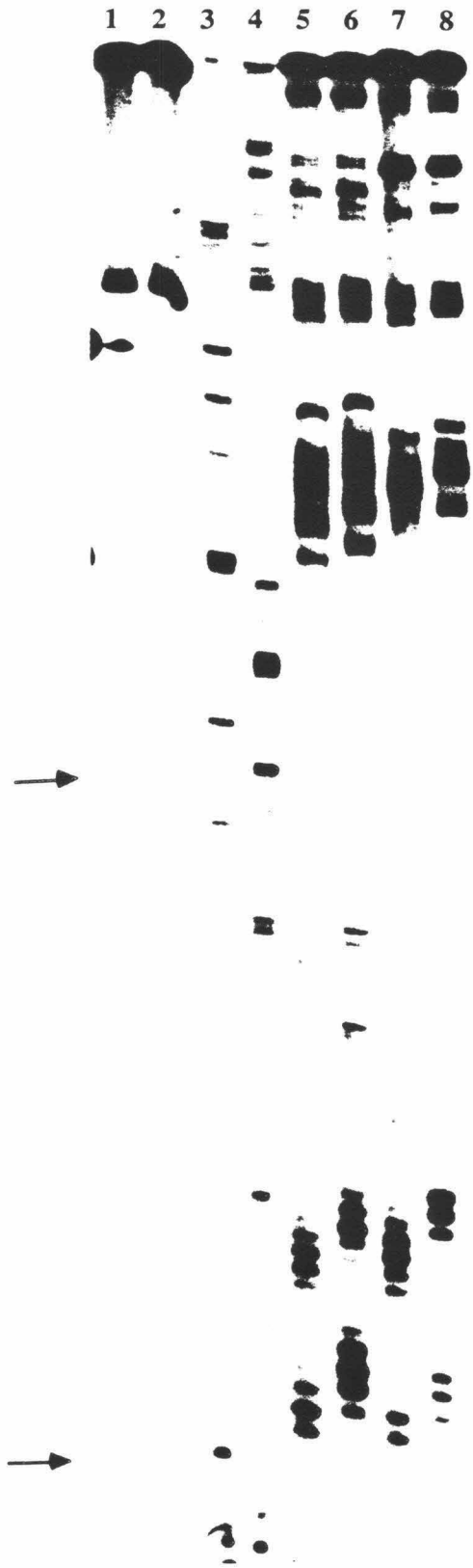


Figure 3.4 Histograms of the DNA cleavage patterns for BEDF·Fe(II) and P7E·Fe(II) on the 167 bp restriction fragment from pBR322. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

Figure 3.5 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 517 bp pBR322 restriction fragment (*Eco*R I-*Rsa* I). The reactions were 100 μM in total DNA base pairs (final concentration). The iron chelates of the compounds were allowed to equilibrate with the DNA for one hour at 25°C. Cleavage was initiated by the addition of DTT, to a final concentration of 5 mM, and allowed to proceed one hour at 25°C.

Odd numbered lanes are 5′-endlabeled DNA. Even numbered lanes are 3′-endlabeled DNA. Lanes 1 and 2 are intact DNA control lanes after incubation under reaction conditions. Lanes 3 and 4 are chemical sequencing lanes containing the products of Maxam-Gilbert G reactions. Lanes 5 and 6 are DNA cleavage reactions with BEDF·Fe(II) at 2.5 μM; Lanes 7 and 8 are DNA cleavage reactions with P7E·Fe(II) at 6.6 μM.


1 2 3 4 5 6 7 8

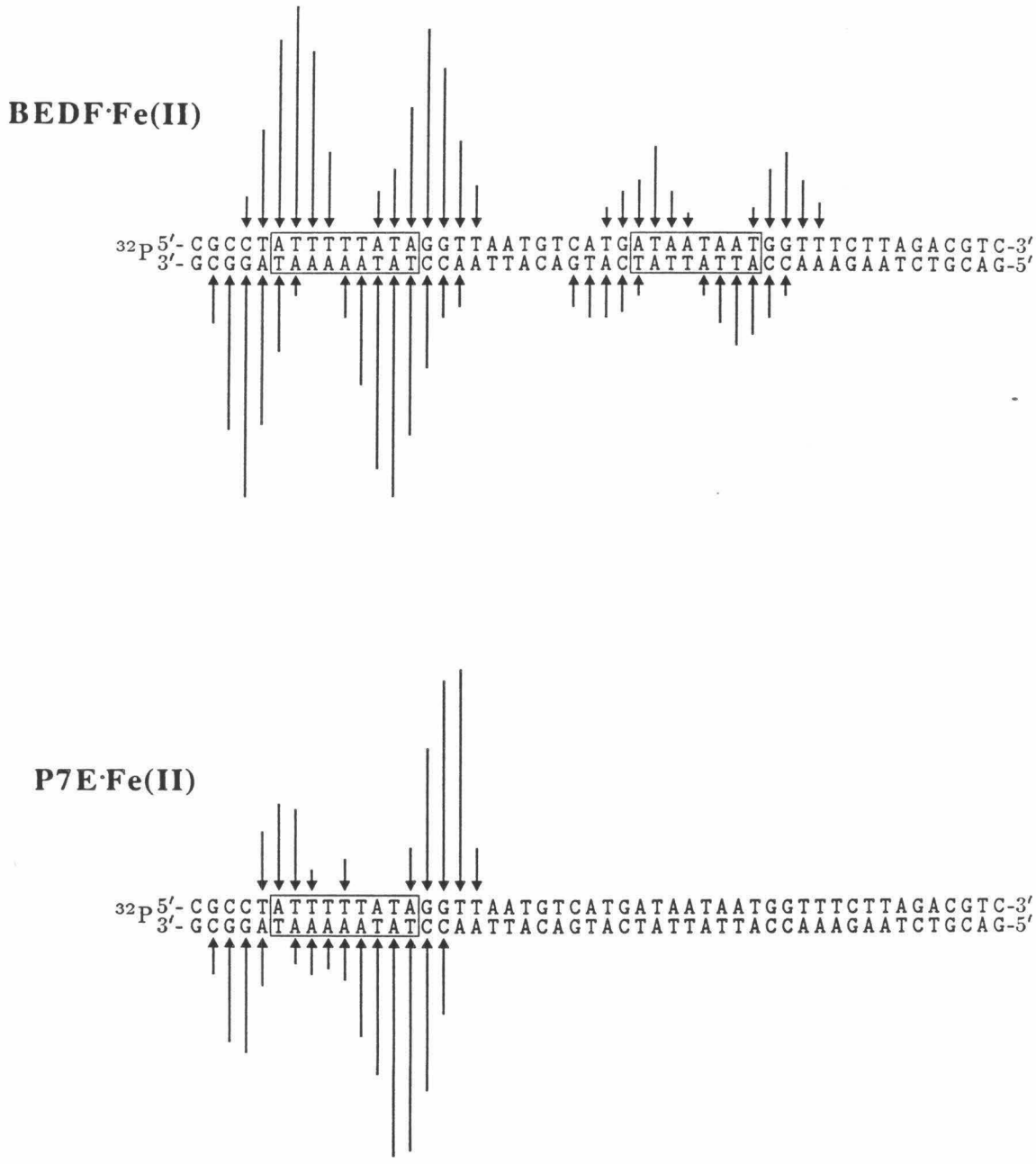


Figure 3.5 Histograms of the DNA cleavage patterns for BEDF·Fe(II) and P7E·Fe(II) on the 517bp restriction fragment from pBR322. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

Figure 3.6 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 517 bp pBR322 restriction fragment (*Eco*R I-*Rsa* I) utilizing optimized BEDF·Fe(II) cleaving conditions. The reactions were 100 μM in total DNA base pairs (final concentration). The iron chelates of the compounds were allowed to equilibrate with the DNA for two hours at 37°C. Cleavage was initiated by the addition of DTT, to a final concentration of 5 mM, and allowed to proceed two hours at 37°C.

Odd numbered lanes are 5′-endlabeled DNA. Even numbered lanes are 3′-endlabeled DNA. Lanes 1 and 2 are intact DNA control lanes after incubation under reaction conditions. Lanes 3 and 4 are chemical sequencing lanes containing the products of Maxam-Gilbert G reactions. Lanes 5 and 6 are DNA cleavage reactions with BEDF·Fe(II) at 1.5 μM; Lanes 7 and 8 are DNA cleavage reactions with BEDF·Fe(II) at 0.5 μM.


Control
G
BEDF·Fe(II)
BEDF·Fe(II)
5' 3' 5' 3' 5' 3' 5' 3'
1 2 3 4 5 6 7 8

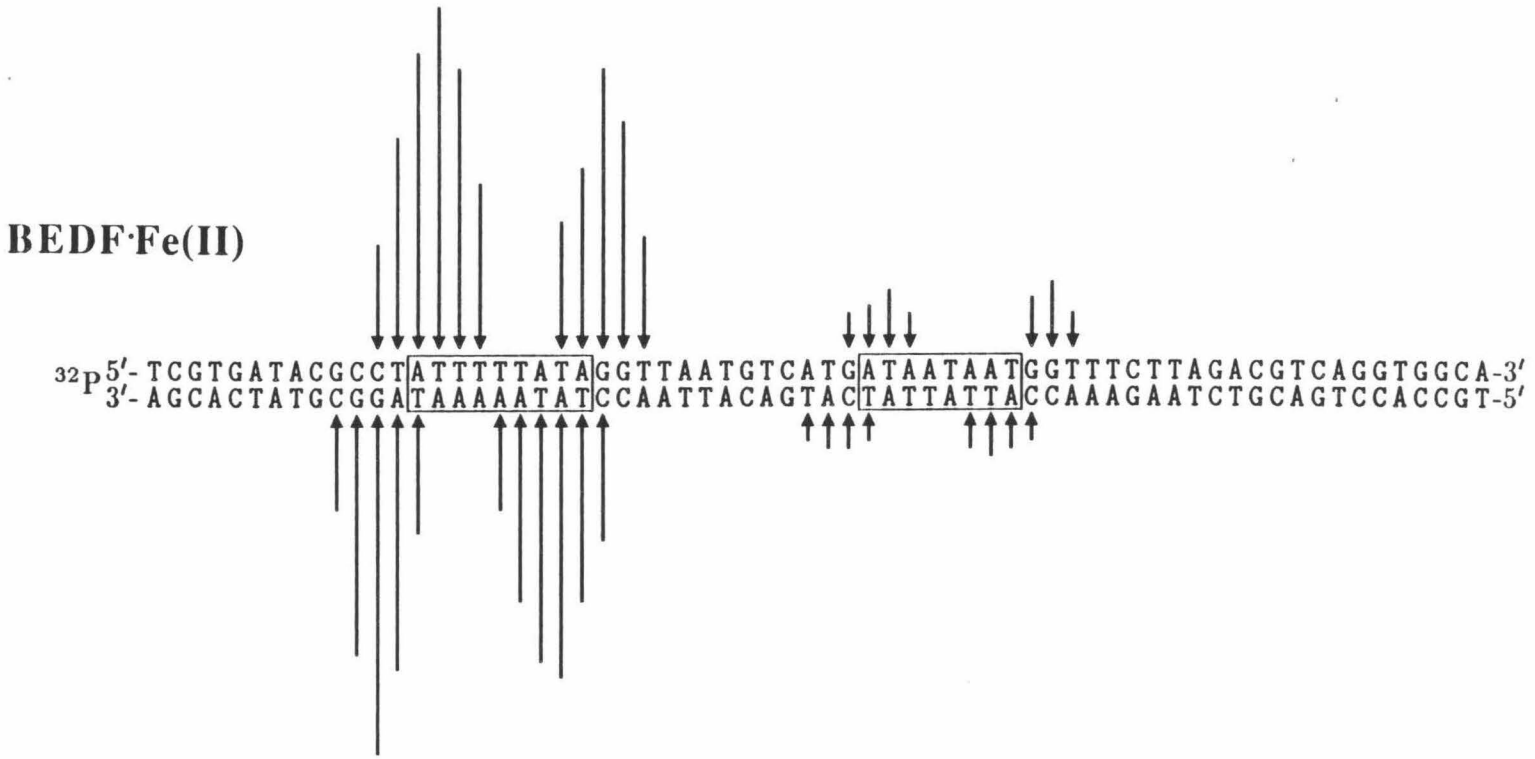


Figure 3.7 Histograms of the DNA cleavage patterns for BEDF·Fe(II) under optimized cleavage conditions on the 517 bp restriction fragment from pBR322. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

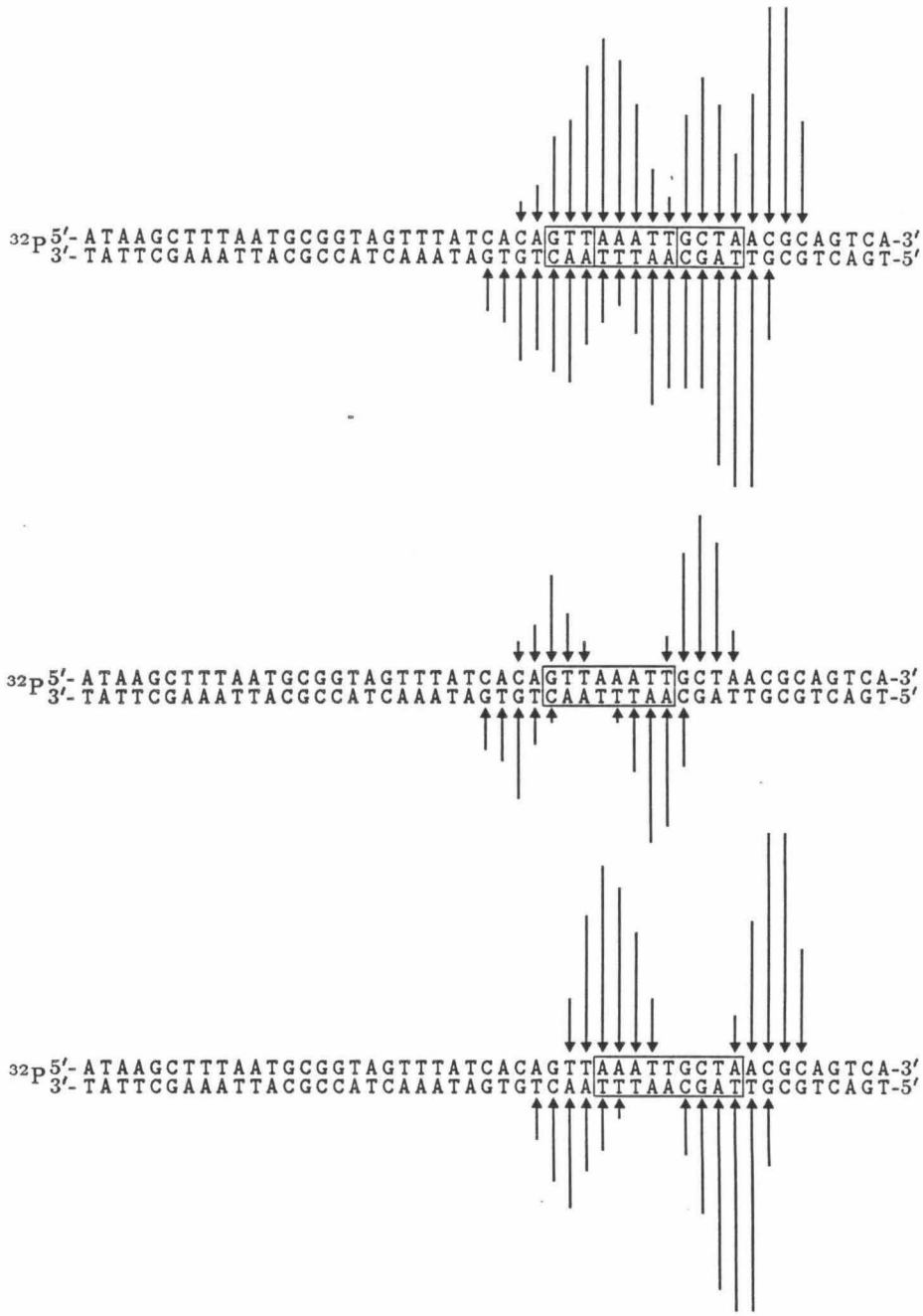


Fig. 3.8 (above) Histogram of the cleavage pattern for BEDF·Fe(II) on the 167bp fragment (*Eco*R I- *Rsa* I, 1-167) from pBR322. (center, below) Histograms of the cleavage components used to assign each of the overlapping binding sites.

Subsequent work with BEDF·Fe(II) found that equilibration at 37°C for 2 hours followed by reaction at 37°C for 2 hours resulted in increased cleavage efficiency. Comparison of BEDF·Fe(II) cleavage under both the original and optimized reaction conditions on the 517 bp fragment results in the same binding sites and sizes (fig. 3.6 and fig. 3.8). This is consistent with the optimized conditions simply increasing the percentage of BEDF·Fe(II) that binds productively without changing either site selection or relative binding affinities.

Binding Site Size. From the studies with the homologous series of polypyrroles a relationship between the number of carboxamides and binding site size was determined. This rule states that a poly-*N*-methylpyrrolecarboxamide molecule containing n amides will bind $n + 1$ base pairs of DNA. Both BEDF·Fe(II) and P7E·Fe(II) contain 8 amides and therefore would be predicted to have binding site sizes of 9 base pairs. Each of the binding sites for P7E·Fe(II) are nine base pairs as calculated. BEDF·Fe(II) was assigned two 9 base pair binding sites and two 8 base pair binding sites (table 3.1). Importantly, no five base pair binding sites were observed as in the case of BED·Fe(II). This demonstrates that the incorporation of a rigid fumaramide linking element into an ED·Fe(II) dimer produces a molecule which binds the DNA exclusively in a dimeric mode.

Interestingly, both eight and nine base pair binding sites for BEDF·Fe(II) were observed. This may be due to the inability to assign the binding site boundaries exactly based on the asymmetric cleavage model described earlier. The element linking the EDTA to the polypyrrole subunit is different than in the

Table 3.1: Observed BEDF·Fe(II) and P7E·Fe(II) binding sites on the 517 and 167 base pair pBR322 restriction fragments

Oligo-peptide	Site(5′–3′)	Site size (base pairs)
	167 bp Fragment	
BEDF	G T T A A A T T	8
BEDF	A A A T T G C T A	9
P7E	A G T T A A A T T	9
	517 bp Fragment	
BEDF	A T T T T T A T A	9
BEDF	A T A A T A A T	8
P7E	A T T T T T A T A	9

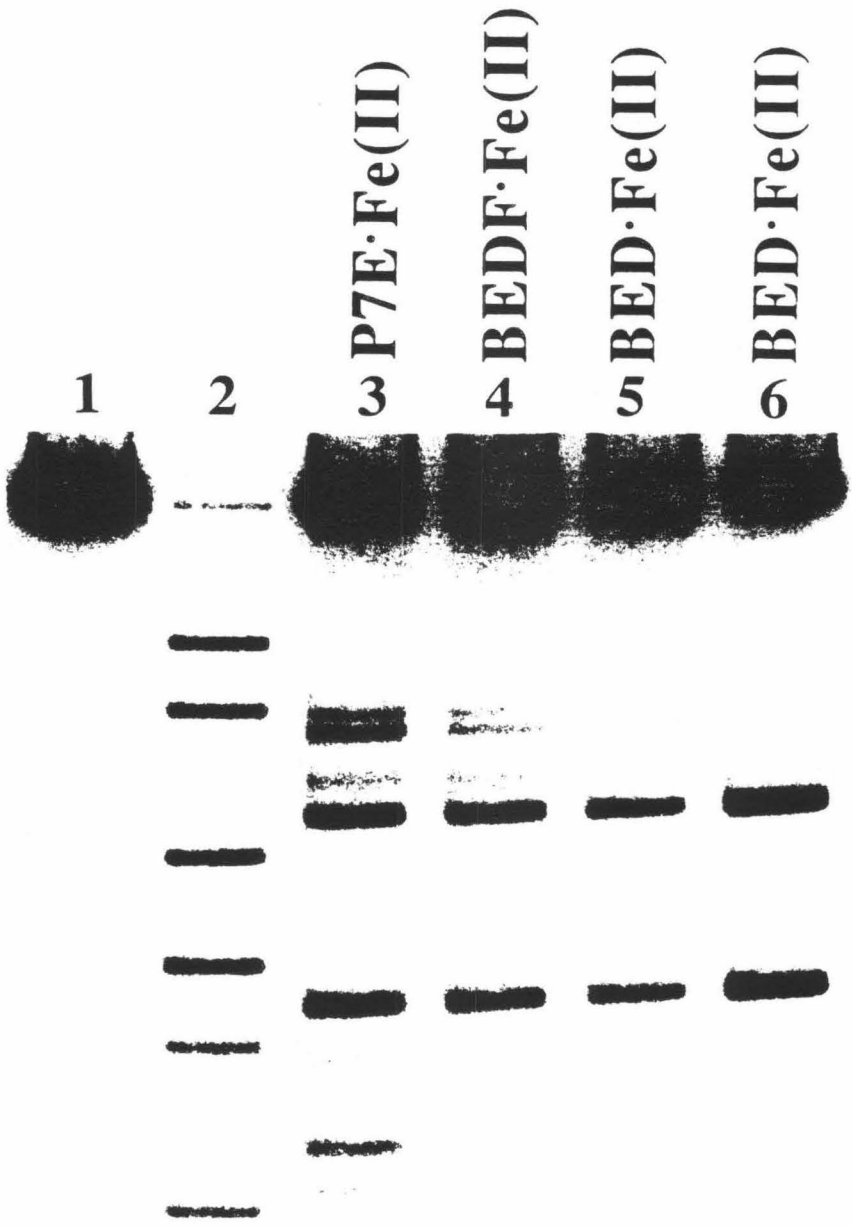
homologous series of compounds containing 3–9 *N*-methylpyrrolecarboxamide units. This results in subtle differences in the cleavage patterns. Binding site assignment in the homologous series of compounds is simplified because it is possible to look for incremental changes between two consecutive members of the series. A second explanation for the eight and nine base pair binding site sizes is that there are different conformations for the fumaramide linking element. This could result in two different binding site sizes. The fumaramide linker was chosen because of its ability to adopt an *s-trans, s-cis* conformation similar in size and shape to an *N*-methylpyrrolecarboxamide, but this may not be the lowest energy conformation for fumaramide. If the fumaramide linking element adopted the *s-trans, s-trans* conformation, CPK models indicate that the polypyrrole compound would have a sharp kink which could decrease the binding site size by one base pair.

DNA Recognition. DNA is recognized by BEDF·Fe(II) and P7E·Fe(II) similarly as would be expected for two octaamide pyrrole compounds. Each of the affinity cleaving compounds recognizes the same region strongly on each of the restriction fragments; however, the weaker binding sites are different in each case. P7E·Fe(II) contains a string of seven consecutive *N*-methylpyrrolecarboxamide residues causing all of the DNA recognition elements to have the same orientation (fig. 3.2). BEDF·Fe(II) is a tail-to-tail dimer on fumarate. This causes the three pyrroles of one subunit to have an orientation opposite that of the other subunit producing a a molecule with an approximate 2-fold symmetry. Possibly, the differences in DNA recognition result from the differences in pyrrole orientation between the two molecules. However, as discussed earlier, the fumaramide linking element also has other conformations available to it and these alternate conformations may be influencing binding site selection.

Large DNA. The high resolution denaturing gel studies show a difference in DNA recognition by P7E·Fe(II) and BEDF·Fe(II). The amount of information available from the experiments on restriction fragments is limited by the small number of unique DNA sequences which they contain. In order to better compare the specificities of these two molecules, double strand cleavage experiments were carried out on intact linearized pBR322. BED·Fe(II) was included in these studies to compare the complete series of octaamide compounds. Cleavage reactions were carried out as described earlier and final concentration, conditions and reaction times are given in the gel figure legends.

Figure 3.10 Autoradiogram of a 1% non-denaturing agarose gel with ^{32}P-endlabeled pBR322 labeled at both ends of the *Sty* I site. The reactions were 100 μM in total DNA base pairs (final concentration). The iron chelates of the compounds were allowed to equilibrate with the DNA for one hour at 65°C. Cleavage was initiated by the addition of DTT to a final concentration of 5 mM and allowed to proceed for 2 hours at 37°C.

Lane 1 is an intact DNA control lane incubated under reaction conditions. Lane 2 contains the molecular weight marker standards. Lane 3 is a DNA cleavage reaction with P7E·Fe(II) at 0.8 μM; Lane 4 is a DNA cleavage reaction with BEDF·Fe(II) at 0.2 μM; Lane 5 is a DNA cleavage reaction with BED·Fe(II) at 0.2 μM; Lane 6 is a DNA cleavage reaction with BED·Fe(II) at 0.4 μM.


1
2
3
P7E·Fe(II)
4
BEDF·Fe(II)
5
BED·Fe(II)
6
BED·Fe(II)

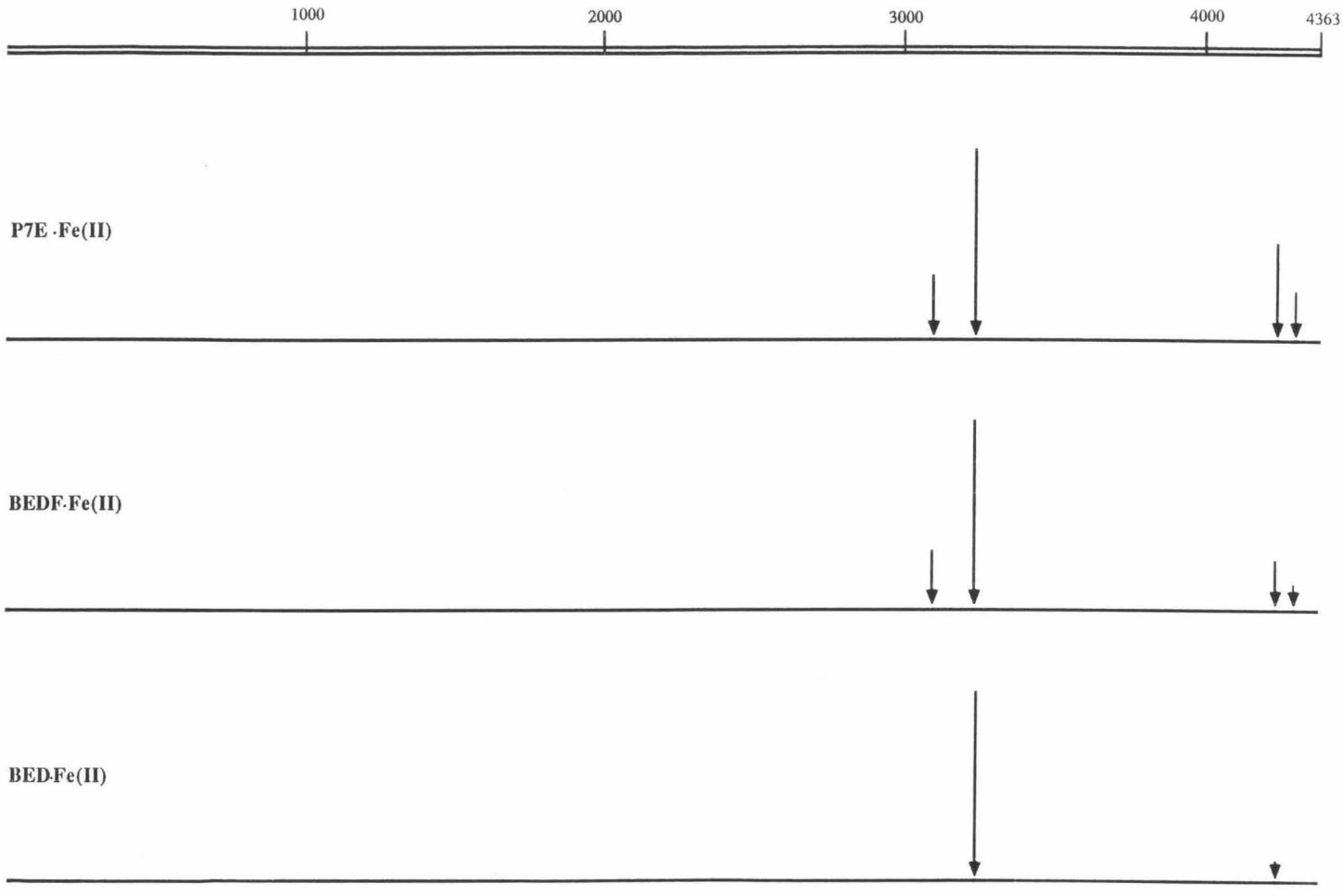


Figure 3.11 Histogram analysis of the double strand cleavage of pBR322 by BEDF·Fe(II), P7E·Fe(II), and BED·Fe(II) from figure 3.10. Arrows represent the amount of cleavage produced at each site.

Each of the molecules produces a discrete banding pattern. These patterns were quantitated and converted into histogram form (fig. 3.11). The binding site assignments were based on the assignments for P7E·Fe(II) described earlier. Two regions on pBR322 are cleaved by all of the octaamide compounds. One site contains the longest run of A·T DNA on pBR322 with 15 consecutive A·T residues (3235 on pBR322). The other is a sequence containing a stretch of nine consecutive A·T residues with only one base interupting the run of polypyrimidines (4245 on pBR322). The two additional cleavage sites observed with BEDF·Fe(II) and P7E·Fe(II) include the longest polypyrimidine run of consecutive A·T base pairs on pBR322, seven at 3080, and the second longest run of A·T DNA on pBR322, ten base pairs at 4325. Importantly, the sites on the 167 bp fragment do not show up in this assay. This means that while the sites observed on the 167 bp fragment are strong with respect to other sites on that fragment, they are weak compared to these sites.

Table 3.2: Binding Sites for the Octaamide Compounds on pBR322

Site	Sequence	P7E	BEDF	BED
3080	GGTGGTTTTTTTGTTTGCAA	M	M	
3235	GATCCTTTTAAATTAAAAAT	S	S	S
4245	TGTATTTAGAAAAATAAACA	M	M	W
4325	TTAACCTATAAAAATAGGCG	M	W	

S strong; M moderate; W weak.

The strong similarity in the cleavage patterns between P7E·Fe(II) and BEDF·Fe(II) demonstrates that the *N*-methylpyrrolecarboxamide can be efficiently replaced by a structurally simpler linking element to produce a molecule

which maintains the sequence specificity of a poly-*N*-methylpyrrolecarboxamide. Because the same sites are cleaved by both molecules, it suggests that amide orientation is not important for the recognition of the highest affinity sites. In these experiments, P7E·Fe(II) was used at 0.8 μM and both BED·Fe(II) and BEDF·Fe(II) were used at 0.2 μM. These concentrations were chosen because they produce approximately the same amount of cleavage at the highest affinity site. The higher cleavage efficiency observed with BED·Fe(II) and BEDF·Fe(II) may be due to a higher affinity for these compounds to the DNA. The higher affinity could arise from the linking elements permitting the poly-*N*-methylpyrrolecarboxamide subunits to more closely match the DNA recognition pattern. An alternative explanation for the higher cleavage efficiency is that the incorporation of two EDTA·Fe(II) cleaving functionalities into the dimers increases the likelihood of double strand cleavage in a single binding event and thereby makes double strand cleavage more efficient.

Specificity. These results show that BED·Fe(II) is much more specific than either P7E·Fe(II) or BEDF·Fe(II) in the double strand cleavage assay and that specificity decreases in the series of octaamide compounds BED·Fe(II) $\gg$ BEDF·Fe(II) $\geq$ P7E·Fe(II). This was an unexpected result because the high resolution gel studies with BED·Fe(II) showed both monomeric and dimeric binding; this should produce lower sequence specificity than dimeric binding alone. Mapping the cleavage sites for BED·Fe(II) from the double strand studies shows that none of the sites studied in the high resolution work were observed in the double

strand assay and; therefore, they were not BED·Fe(II)'s highest affinity sites. This is consistent with monomeric binding as a weaker binding mode which is not employed by BED·Fe(II) at its highest affinity sites.

BEDF·Fe(II) and P7E·Fe(II) are two octaamides which recognize the same sites on pBR322 with approximately the same affinities while using different poly-*N*-methylpyrrolecarboxamide orientations. BED·Fe(II) is also a poly-*N*-methylpyrrolecarboxamide compound similar in structure to BEDF·Fe(II); however, it is more specific in its cleavage of pBR322 than either of the other octaamide compounds. This increase in specificity may result from the flexibility that the heptanedioic acid linking element provides BED·Fe(II).

P7E·Fe(II) and BEDF·Fe(II) are rigid molecules which require simultaneous recognition of the DNA by all of their amides. Furthermore, all of the amides are rigidly set at fixed distances. Therefore, BEDF·Fe(II) and P7E·Fe(II) have very little freedom in the way they align their amides with the DNA base pairs. Because heptanedioic acid is a long flexible tether and monomeric binding has been documented for BED·Fe(II), it can be assumed that heptanedioic acid is a long enough linker to allow independent recognition of the DNA by the two tripyrrole subunits. This provides BED·Fe(II) with a large degree of freedom in the spacing between the two polypyrrole subunits. Biophysical studies have shown that polypyrrole compounds alter DNA structure and have provided evidence of cooperative binding.[33,34,51,99] BED·Fe(II) has the flexibility for one subunit to bind and alter the DNA conformation and allow the second subunit

to make adjustments and bind cooperatively. This binding mechanism superimposes the effects of cooperativity and anti-cooperativity on the normal sequence specificity of the tri-*N*-methylpyrrolecarboxamide monomers, amplifying the differences in binding energy for different sites. P7E·Fe(II) and BEDF·Fe(II) with their rigid structures and fixed amide distances probably do not have the flexibility required for cooperative DNA recognition. The increased specificity in the series of the poly-*N*-methylpyrrolecarboxamide octaamides could result from the flexible linking element in BED·Fe(II) allowing its two subunits to bind to DNA independently with allostery while the rigid structure of P7E·Fe(II) and BEDF·Fe(II) precludes this type of interaction.

This result presents a new strategy for DNA recognition: oligomerization of poly-*N*-methylpyrrolecarboxamide subunits with flexible tethers. This approach would allow the sequence specificity of the individual subunits to be enhanced by cooperative binding interactions. By utilizing allosteric binding, this strategy in effect makes the specificity of the whole greater than the sum of the parts.

Recent studies in our group have focused on amino acid linking elements for two di-*N*-methylpyrrolecarboxamide subunits.[104] This work found that β-alanine was a more effective tether than either glycine or γ-amino butyric acid. Incorporation of β-alanine into the dimers resulted in solely dimeric binding demonstrating that a rigid linker is not required for dimeric binding. These dimers also showed very efficient DNA recognition as evidenced by efficient DNA cleavage.

$(P4)_3E\cdot Fe(II)$. The knowledge that β-alanine is a flexible amino acid linking element promoting dimeric binding and efficient DNA recognition makes possible the design of an oligomer of *N*-methylpyrrolecarboxamides capable of recognizing very long stretches of A·T containing DNA (fig. 3.12). This approach resulted in a new sequence-specific DNA binding molecule, tris(tetra-*N*-methylpyrrolecarboxamide)bis(β-alaninyl)-EDTA **36** ($(P4)_3E$) (fig. 3.13). This pentadecaamide contains three tetra-*N*-methylpyrrolecarboxamide subunits coupled by β-alanine and according to the $n + 1$ rule should bind 16 base pairs of DNA. Attachment of EDTA to the amino terminus of the trimer allows the use of the affinity cleaving method to determine the binding site size and sequence through analysis of cleavage patterns on ^{32}P-endlabeled DNA restriction fragments. As a point of reference, the frequency at which a 16 A·T base pair stretch occurs is equivalent to that of a unique 8 base pair site (table 2.6). Consequently, molecules of this type may have the specificity needed to operate on very large DNA.

Synthesis. The synthesis of $(P4)_3E$ employed a highly convergent synthetic route utilizing tetra-*N*-methylpyrrolecarboxamide nitro ester **37** as the common intermediate for all three tetra-*N*-methylpyrrolecarboxamide subunits (scheme 3.2). As described previously, tetra-*N*-methylpyrrolecarboxamide nitro ester **37** can be converted to tetra-*N*-methylpyrrolecarboxamide nitro dimethyl amine **38**. Reduction of the terminal nitro group followed by acylation with the imidazolide of *t*-boc β-alanine **39** afforded tetra-*N*-methylpyrrolecarboxamide

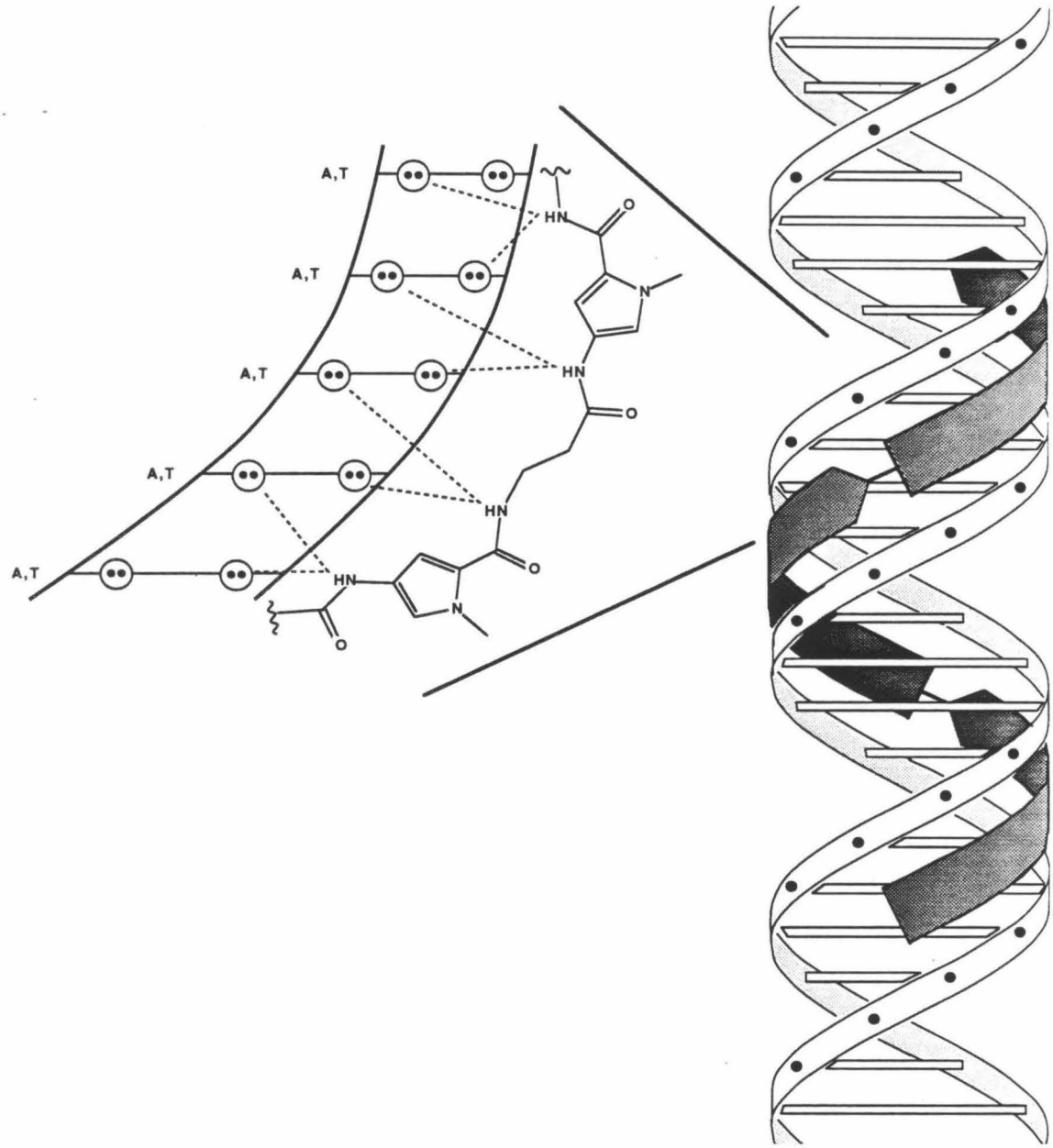


Figure 3.12 Model for β-alanine linked trimer recognition of double helical DNA.

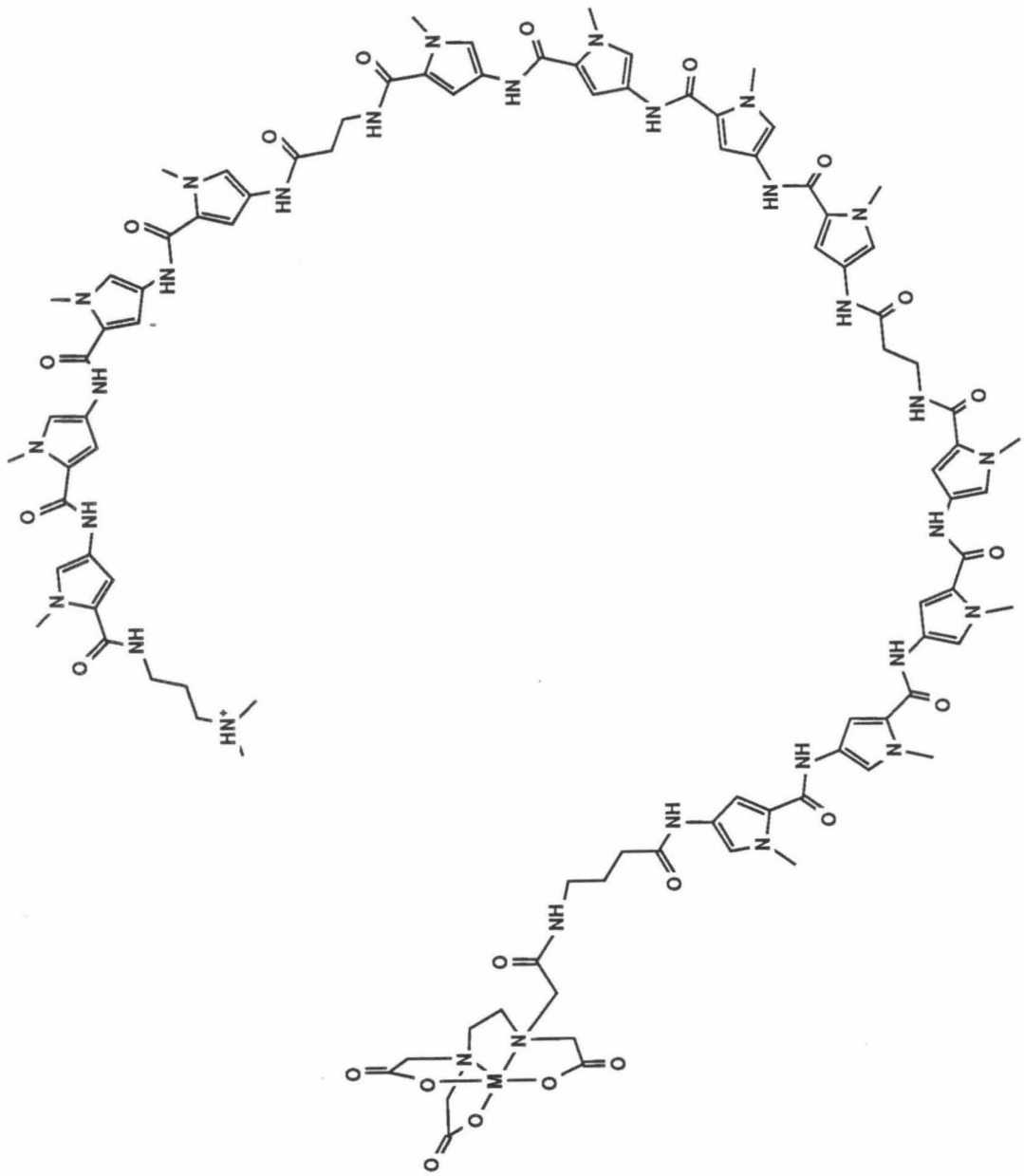
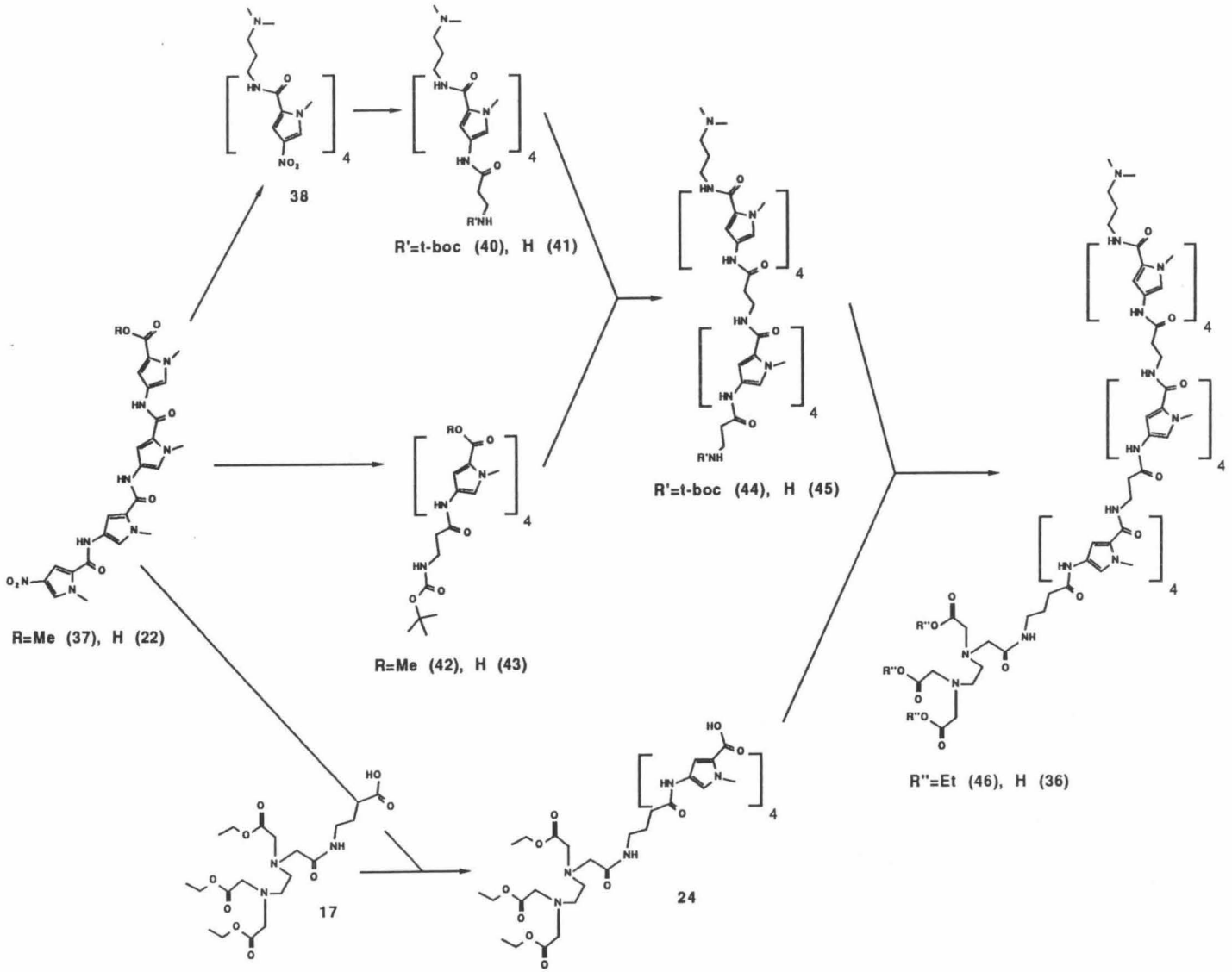


Figure 3.13 $(P4)_3E{\cdot}M$

dimethyl amino *t*-boc β-alanine **40**. The primary amine was then deprotected with TFA in CH_2Cl_2 to yield tetra-*N*-methylpyrrolecarboxamide dimethyl amino β-alanine **41**. Reduction of the terminal nitro group of tetra-*N*-methylpyrrolecarboxamide nitro ester **37** followed by acylation with the imidazolide of *t*-boc β-alanine produced tetra-*N*-methylpyrrolecarboxamide methyl ester *t*-boc β-alanine **42**. Hydrolysis of the methyl ester with sodium hydroxide afforded tetra-*N*-methylpyrrolecarboxamide acid *t*-boc β-alanine **43**. Condensation of tetra-*N*-methylpyrrolecarboxamide dimethyl amino β-alanine **41** and tetra-*N*-methylpyrrolecarboxamide acid *t*-boc β-alanine **43** using dicyclohexylcarbodiimide and *N*-hydroxybenzotriazole produced the *t*-boc protected tetra-*N*-methylpyrrolecarboxamide β-alanine dimer **44** which was deprotected with TFA in CH_2Cl_2 to afford tetra-*N*-methylpyrrolecarboxamide β-alanine dimer **45**. The tetra-*N*-methylpyrrolecarboxamide acid EDTA triethyl ester **24** is available from tetra-*N*-methylpyrrolecarboxamide nitro ester as previously described. Coupling **24** with the tetra-*N*-methylpyrrolecarboxamide dimer **45** produced the triethyl ester of the final compound, $(P4)_3$E-3E **46**. Hydrolysis of the ethyl esters with lithium hydroxide afforded the final compound, $(P4)_3$E **36**, a molecule with a molecular weight of 2154.

DNA recognition by $(P4)_3$E was examined by both MPE·Fe(II) footprinting and affinity cleaving on a ^{32}P-endlabeled restriction fragments from plasmid SV-CAT: a 205 bp fragment, *Xba* I-*Hin*d III (5174-129 on SV40 and six base pairs of the pUC 18 polylinker region). MPE·Fe(II) footprinting reactions were



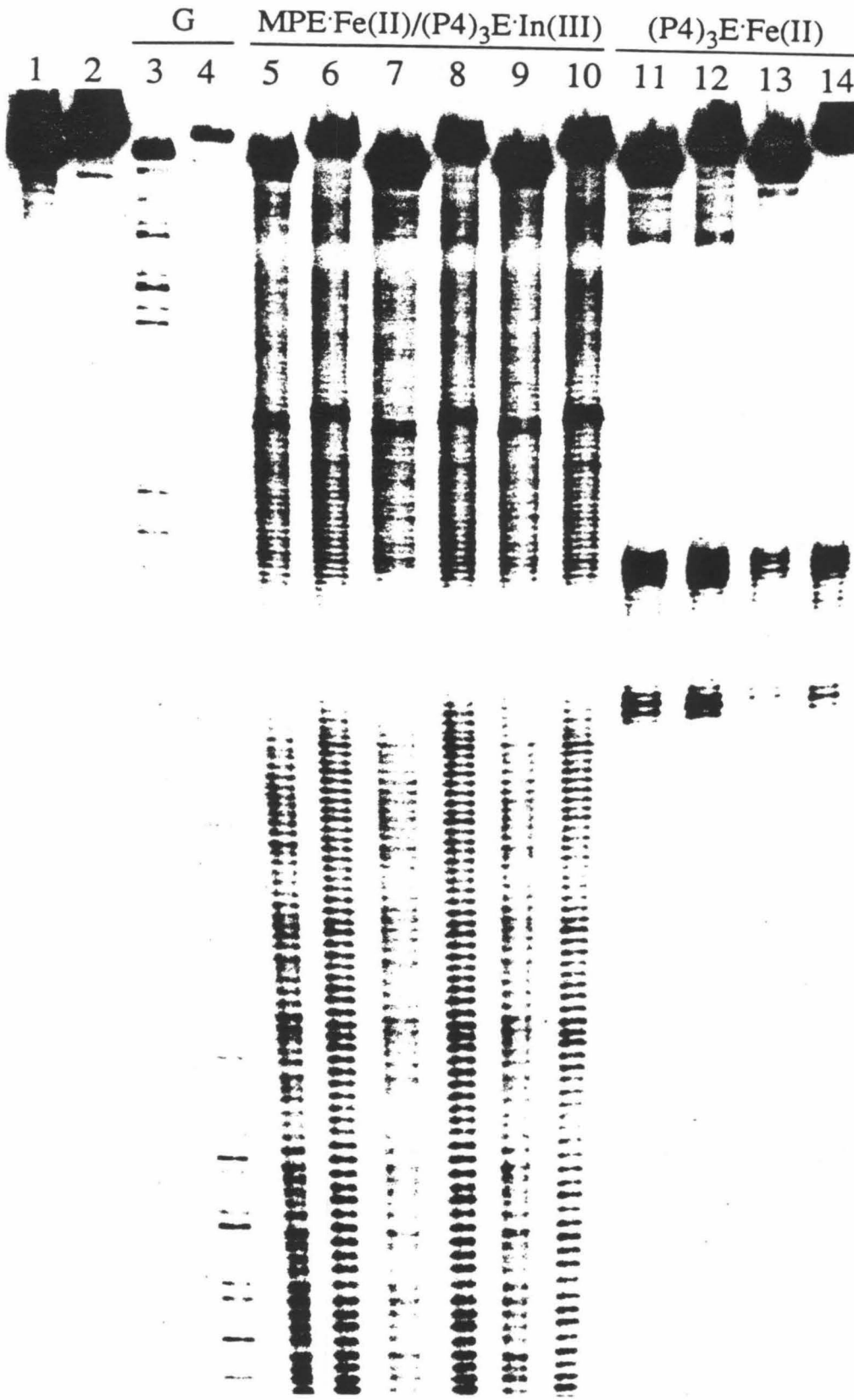
Scheme 3.2 Synthesis of $(P4)_3E$

carried out as described earlier using the indium chelate of $(P4)_3E$. The indium chelate was formed by mixing $(P4)_3E$ with 2 equivalents of indium(III) chloride. Indium has been shown to effectively sequester the EDTA chelator and prevent affinity cleaving under redox conditions.[104] The affinity cleaving reactions were carried out as described earlier. The final reaction conditions and concentrations for both the footprinting and affinity cleaving reactions are given in the figure legends.

Densitometric analysis of the footprinting data for $(P4)_3E{\cdot}In(III)$ is consistent with the occupation a 16 bp stretch of A·T DNA (fig. 3.14, 3.15). Analysis of the affinity cleaving data (fig. 3.14, 3.15) reveals two major cleavage loci flanking a 16 base pair A·T site. According to the $n + 1$ rule, a functional pentadecaamide should recognize 16 base pairs of DNA. The observation of only two cleavage loci with affinity cleaving and the footprinting data demonstrate that $(P4)_3E$ is binding to the DNA with a single binding mode. These patterns define a 16 bp region of A·T DNA which is assigned as the binding site of $(P4)_3E$. The observation of *exclusive* 16 base pair binding in the absence of possible monomeric (6 bp) or dimeric (11 bp) sites demonstrates that the β-alanine linker is an efficient replacement for the *N*-methylpyrrolecarboxamide and verifies that a rigid tether is not required to assure simultaneous binding of all subunits. *N*-methylpyrrolecarboxamide compounds are known to bind B-form DNA in solution.[17, 18, 7, 20–22] If the $(P4)_3E$-DNA complex is in the B-form, then this synthetic organic molecule is recognizing a turn and a half of the DNA helix.

Figure 3.14 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 205 bp SV-CAT restriction fragment (*Xba* I-*Hin*d III). The reactions were 100 μM in total DNA base pairs (final concentration). For the affinity cleaving reactions, the $(P4)_3E{\cdot}Fe(II)$ was allowed to equilibrate with the DNA for one hour at 65°C. Cleavage was initiated by the addition of DTT to a final concentration of 3 mM and allowed to proceed 2 hours at 37°C. For the footprinting reactions, the $(P4)_3E{\cdot}In(III)$ was allowed to equilibrate with the DNA for one hour at 65°C. Cleavage was initiated by the addition of MPE·Fe(II) at 3 μM and DTT to a final concentration of 3 mM and was allowed to proceed 15 min. at 37°C.

Odd numbered lanes are 5′-labeled DNA. Even numbered lanes are 3′-labeled DNA. Lanes 1 and 2 are intact DNA control lanes incubated under reaction conditions. Lanes 3 and 4 are chemical sequencing lanes containing the products of Maxam-Gilbert G reactions. Lanes 5 and 6 are MPE·Fe(II) DNA cleavage reactions; Lanes 7 and 8 are MPE·Fe(II) DNA cleavage reactions with $(P4)_3E{\cdot}In(III)$ at 1 μM; Lanes 9 and 10 are MPE·Fe(II) DNA cleavage reactions with $(P4)_3E{\cdot}In(III)$ at 0.3 μM; Lanes 11 and 12 are DNA cleavage reactions with $(P4)_3E{\cdot}Fe(II)$ at 1 μM; Lanes 13 and 14 are DNA cleavage reactions with $(P4)_3E{\cdot}Fe(II)$ at 0.3 μM.


G
MPE·Fe(II)/(P4)3E·In(III)
(P4)3E·Fe(II)
1 2 3 4 5 6 7 8 9 10 11 12 13 14

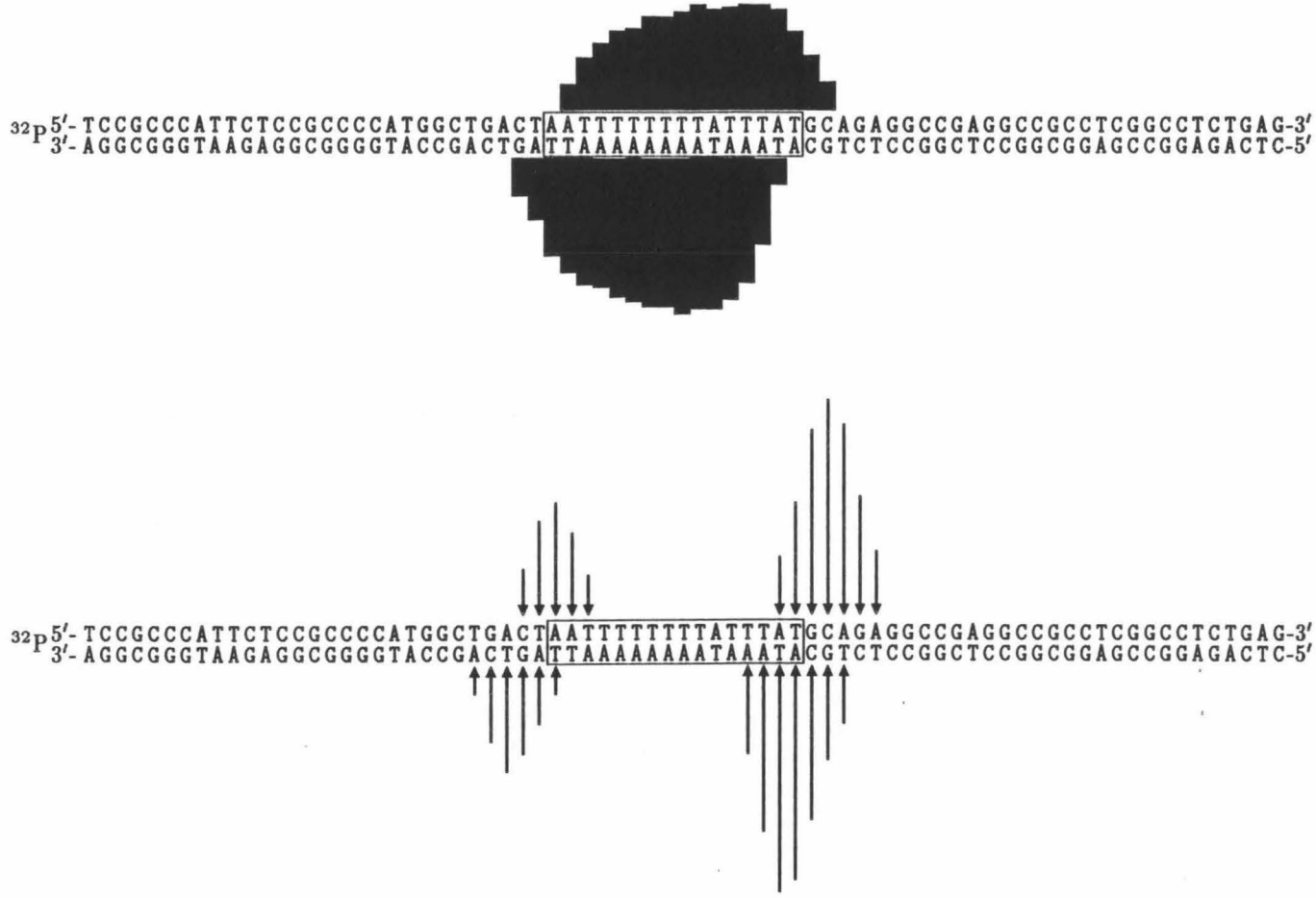


Figure 3.15 (Upper) Histograms of the MPE·Fe(II) DNA cleavage inhibition patterns for $(P4)_3E$·In(III) on the 205 bp SV-CAT fragment. Bars represent the amount of MPE·Fe(II) cleavage inhibition produced by $(P4)_3E$·In(III). Boxes define binding site location and size. (Lower) Histograms of the DNA cleavage patterns for $(P4)_3E$·Fe(II) on the 205 bp SV-CAT fragment. Arrows represent the amount of cleavage resulting in the removal of the indicated base. Boxes define binding site location and size based on the asymmetric cleavage model.

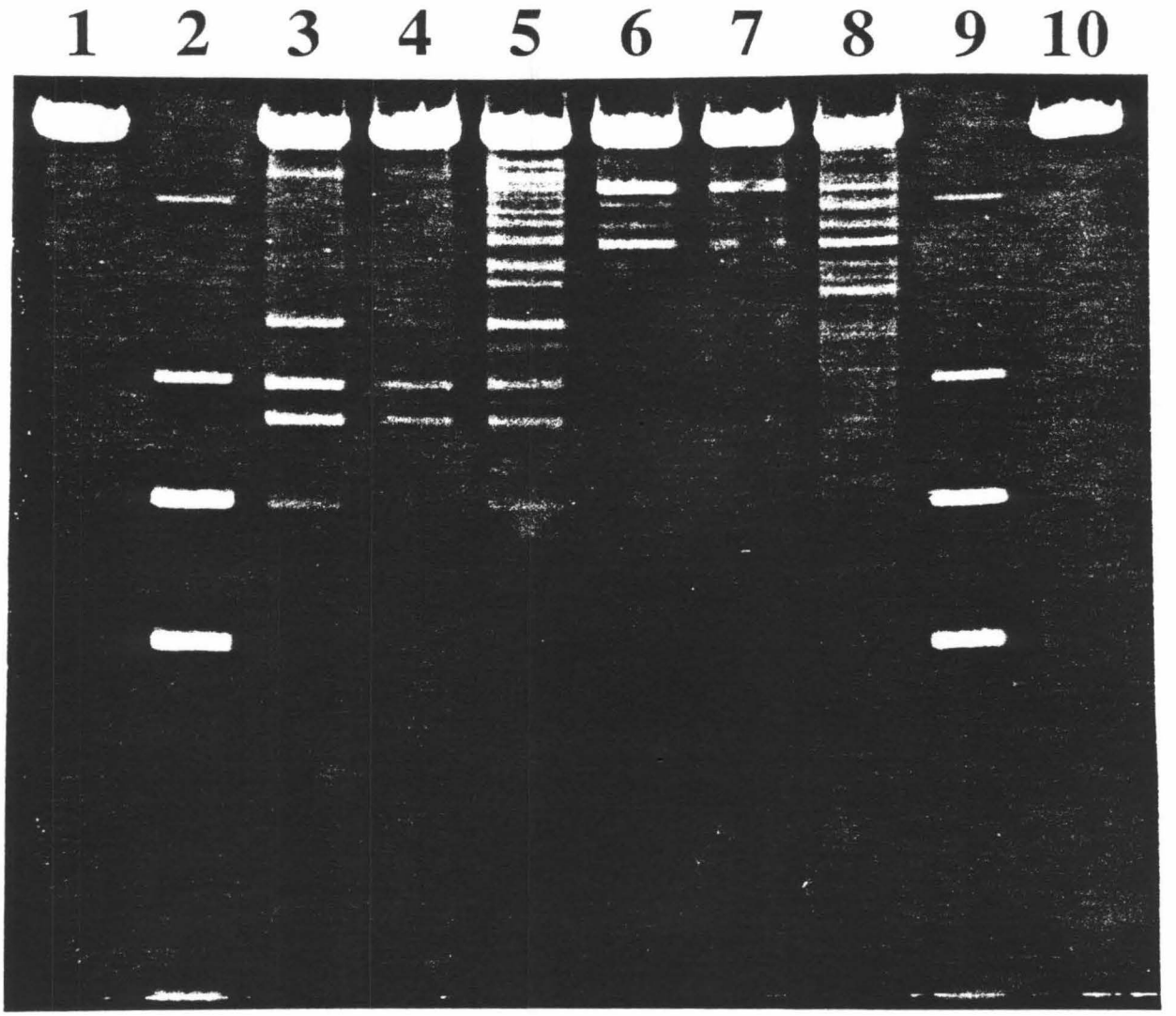
Large DNA. The high resolution studies have demonstrated that $(P4)_3E{\cdot}Fe(II)$ recognizes a 16 bp binding site as predicted by the $n + 1$ rule. No evidence of competing monomeric or dimeric binding modes was observed on this restriction fragment. To determine whether this specificity can be extended to the cleavage of large pieces of DNA, double strand cleavage experiments were carried out on the SV-CAT plasmid. This plasmid was chosen because it includes base pairs 5174-129 of SV40. This region contains the 17 bp run of A·T DNA studied in the high resolution assays and the plasmid itself has another 18 bp run of consecutive A·T DNA ensuring that there are high affinity sites for $(P4)_3E{\cdot}Fe(II)$ on the plasmid. Cleavage reactions were carried out as described earlier for double strand cleavage and final concentration, conditions, and reaction times are given in the gel figure legends.

By inspection, it is apparent that P9E·Fe(II) digestion of SV-CAT (fig. 3.16 lanes 3 and 4, 6 and 7) produces far fewer bands that $(P4)_3E{\cdot}Fe(II)$ (fig. 3.16 lanes 5, 8). This contradicts the expectation that a pentadecaamide should be more specific than a decaamide. It indicates that $(P4)_3E{\cdot}Fe(II)$ is recognizing DNA either with a low sequence specificity or more likely monomeric and dimeric binding modes are interfering with its ability to recognize DNA effectively.

Reaction conditions were varied over a wide range in an effort to overcome the low specificity of double strand cleavage. Pre-equilibration was carried out over a time range of 2-24 hours and a temperature range of 20°C to 90°C. Cleavage reactions were run from 1 to 16 hours at temperatures from 20°C

Figure 3.16 Photograph of a 1% non-denaturing agarose gel with SV-CAT DNA stained with ethidium bromide and visualized by fluorescence. The reactions were 100 μM in total DNA base pairs (final concentration). The $(P4)_3E{\cdot}Fe(II)$ was allowed to equilibrate with the DNA for two hours at 65°C. Cleavage was initiated by the addition of DTT to a final concentration of 5 mM and allowed to proceed for 2 hours at 65°C.

Lanes 1-5 contain SV-CAT linearized with *Bam*H I; Lanes 6-10 contain SV-CAT linearized with *Eco*R V; Lanes 1 and 10 are intact DNA control lanes incubated under reaction conditions; Lanes 2 and 9 contain the molecular weight marker standards; Lanes 3 and 6 are the DNA cleavage reactions with P9E·Fe(II) at 0.1 μM; Lanes 4 and 7 are the DNA cleavage reactions with P9E·Fe(II) at 0.05 μM; Lanes 5 and 8 are the DNA cleavage reactions with $(P4)_3E{\cdot}Fe(II)$ at 0.1 μM.


1
2
3
4
5
6
7
8
9
10

to 65°C. Because organic solvents have been shown to facilitate triple strand formation,[105] reactions were run in 40% ethylene glycol. Finally, different DNA substrates, pBR322 and SV40, were used. None of these modifications had any appreciable effect on the number of bands which were produced in the reaction.

In summary, trimerization of tetra-*N*-methylpyrrolecarboxamide subunits with β-alanine produces a molecule which is capable of recognizing 16 base pairs of A·T DNA, more than a turn and a half of the DNA helix, and, with the attachment of EDTA, is capable of efficient DNA cleavage at these sites. Furthermore, this work demonstrates the generality of the *N*-methylpyrrolecarboxamide as an A·T DNA recognition element and the feasibility of linking multiple DNA-binding subunits together, combining the specificity of the individual subunits to produce a molecule of very high sequence specificity. Because double strand cleavage experiments showed many more cleavage sites than would be expected, the linking element still requires optimization. It may be that a longer linker is needed to allow independent binding of the subunits. A longer linking element may, as in the case of BED·Fe(II), promote more effective recognition of DNA by each of the polypyrrole subunits.

Chapter 4

A Comparison of Three Footprinting Reagents

DNA footprinting is a powerful direct method for determining the binding sites of proteins and small molecules on heterogeneous DNA.[41-44, 46, 93, 106] This technique is based on a sequence specific DNA binding molecule's ability to protect an endlabeled DNA fragment from cleavage at its binding site. The binding site is visualized as a region of diminished cleavage on an autoradiogram of a high resolution denaturing gel (fig. 4.1). The two most commonly used reagents for DNA footprinting are the enzyme DNase I and the small molecule MPE·Fe(II); however, neither reagent is completely sequence neutral. DNase I cleaves at every base, but the amount of cleavage varies greatly from base to base, apparently sensitive to small changes in DNA structure.[107] This property can lead to large regions of the DNA which are very inefficiently cleaved. Regions with low cleavage do not produce a good signal in footprinting experiments, making assignment of the exact positions of the binding sites difficult. MPE·Fe(II) is a synthetic organic molecule with lower sequence specificity than DNase I. MPE·Fe(II) cleaves DNA significantly at every base, but it cleaves preferentially in G·C regions. This results in regions of enhanced and diminished cleavage. This uneven baseline also complicates binding site assignment. Conceptually, then, the ability to cleave DNA in a completely sequence neutral manner would facilitate binding site determination.

Polyamines are thought to recognize DNA sequence neutrally binding as polycations to the polyanionic DNA.[108, 109] Therefore, attachment of a cleaving functionality to a polyamine may produce a DNA-cleaving polyamine useful

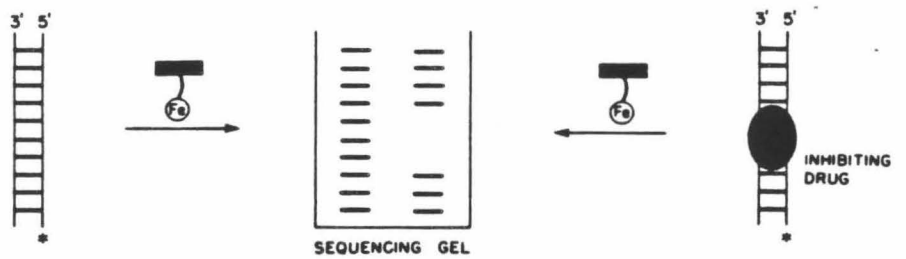


Figure 4.1 Illustration of MPE·Fe(II) footprinting on a high resolution denaturing gel with native and ligand protected segments of double helical DNA.

for DNA footprinting. In addition, such a compound should allow the possible sequence and conformational specificities of polyamines to be studied at the nucleotide level. To explore this approach, EDTA was attached to the polyamine spermine to produce the bifunctional DNA binding/cleaving molecule spermidine-EDTA (SE) (fig. 4.2).

Recent work by Tullius has shown that a system containing EDTA·Fe(II) and ascorbate produces very even DNA cleavage and can be used to footprint proteins.[106] Because EDTA·Fe(II) is negatively charged, it should not bind polyanionic DNA.[110] Cleavage of DNA then requires remote generation of a reactive species which must diffuse to the DNA. MPE·Fe(II) recognizes DNA by

intercalating its aromatic methidium residue between the base pairs. SE·Fe(II) should bind DNA as a polycation electrostatically attracted to the DNA phosphate backbone. To explore the relationship between the way in which the DNA cleaving species is delivered to the DNA and the detection of a small molecule's binding site, a footprinting comparison of three reagents, MPE·Fe(II), SE·Fe(II), and EDTA·Fe(II) was undertaken (fig. 4.2). These compounds are thought to cleave DNA *via* the same reactive species generated from an EDTA·Fe(II) moiety in the presence of O_2 and reducing agent but recognize DNA by different mechanisms.

Synthesis of Spermidine-EDTA. The synthesis of spermidine-EDTA **47** was carried out as shown in scheme 4.1. EDTA-triethyl ester **16** was activated with carbonyl diimidizole and condensed with spermine. The amines were protected with *t*-boc groups using di-*tert*-butyl dicarbonate and the protected spermidine-EDTA **48** was isolated by chromatography on silica gel. The ethyl esters were removed by treatment with 7 equivalents of lithium hydroxide (2.3 eq./ester) followed by *t*-boc removal with triflouroacetic acid. Spermidine-EDTA was isolated as the TFA salt.

DNA Cleavage by MPE·Fe(II), SE·Fe(II), and EDTA·Fe(II). The cleavage patterns produced by partial digestion of a 517 bp (*Eco*R I-*Rsa* I) endlabeled fragment from pBR322 with MPE·Fe(II), SE·Fe(II) and EDTA·Fe(II) were examined by high resolution denaturing gel electrophoresis. The reactions were carried out by equilibrating the cleaving agent (1:50, MPE·Fe(II)/bp or 1:2,

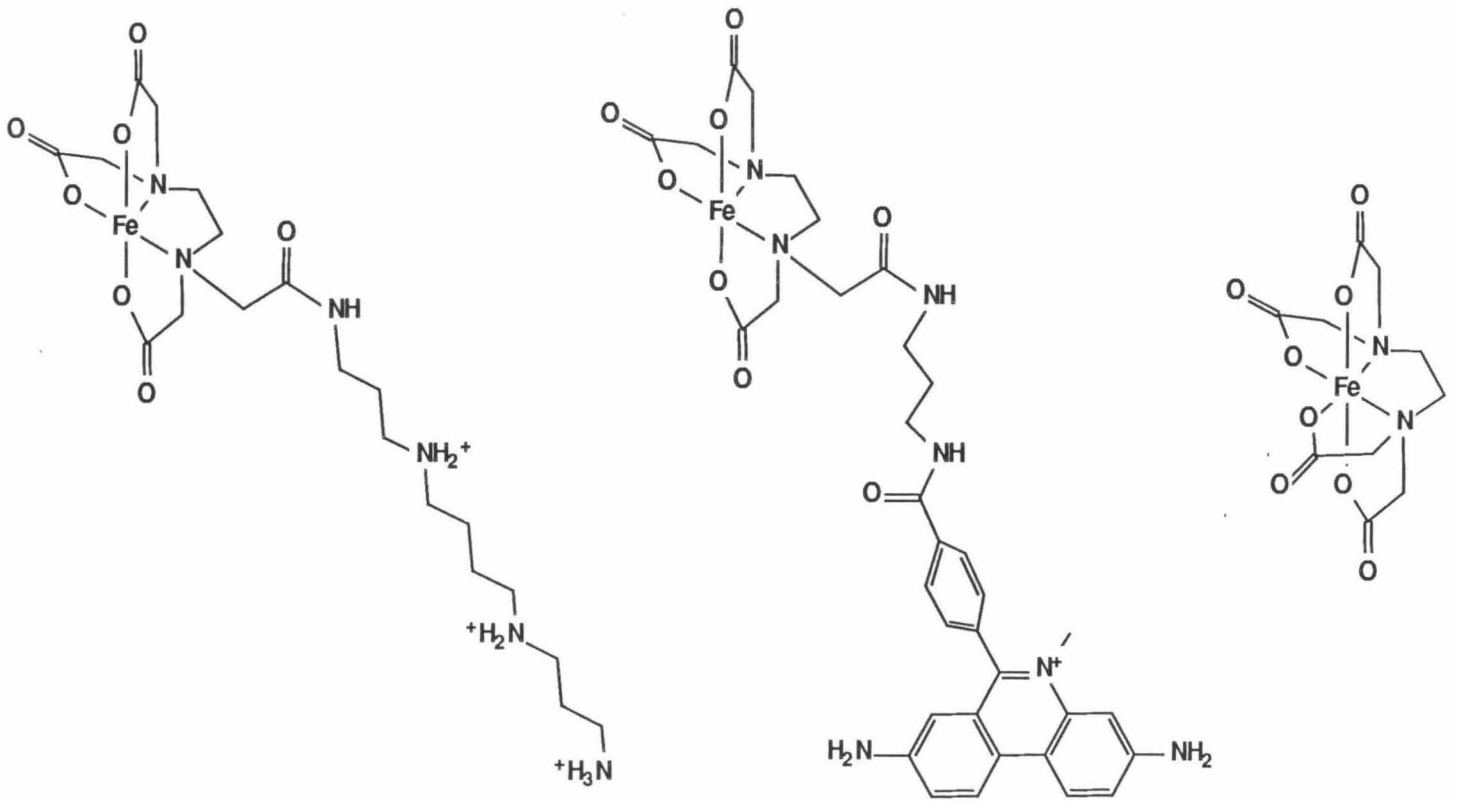
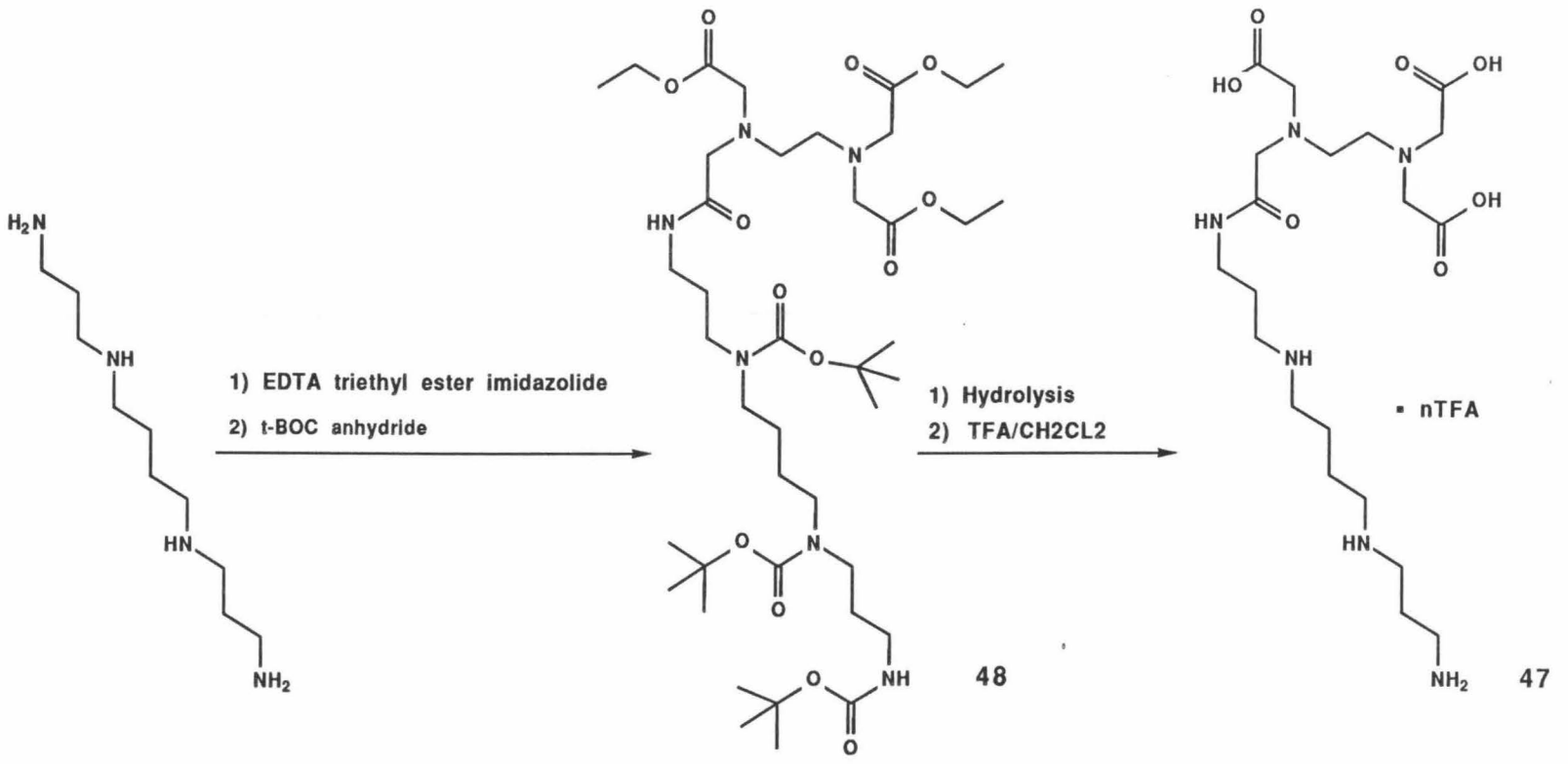


Figure 4.2 Footprinting reagents (left to right) SE·Fe, MPE·Fe, EDTA·Fe.



Scheme 4.1 Synthesis of SE

SE·Fe(II) or EDTA·Fe(II)/bp) with the DNA for 15 minutes at 37°C. Cleavage was initiated by the addition of reducing agent (3 mM DTT or 1 mM ascorbate, final concentration) and the reactions were allowed to continue for 15 minutes at 37°C. The reaction mixtures were frozen, lyophilized, and resuspended in formamide loading buffer. After heat denaturation, they were run on 8% high resolution denaturing polyacrylamide gels. The gels were dried and exposed to photographic film to produce the autoradiograms shown in figures 4.3 and 4.4. The significant features to note in figure 4.3 lanes 3-5, 14-16 and figure 4.4 lanes 3-5, 14-16 are the very even cleavage patterns produced by SE·Fe(II) and by EDTA·Fe(II) with ascorbate and the regions of enhanced and diminished cleavage seen with MPE·Fe(II). These features are quantitated by scanning the gel and plotting distance scanned versus absorbance. The densitometry traces for figure 4.3 lanes 14-16 and 20-11 are shown in figure 4.5. Traces B and C show the very even cleavage pattern seen with SE·Fe(II) and EDTA·Fe(II) respectively. Trace A demonstrates the uneven cleavage pattern seen with MPE·Fe(II). It is also important to note the lack of cleavage seen with EDTA·Fe(II) and DTT (fig. 4.4 lane 5, 8, 11, 16, 19 22).

Footprinting P5. The footprinting reactions were run by equilibrating the DNA with an appropriate concentration of P5 (see fig. 4.3 and 4.4 for final P5 concentrations) for 1 h at 37°C. The cleavage reactions were carried out as described above. The addition of P5 produces regions of diminished cleavage. The assigned P5 binding sites and cleavage inhibition patterns for both concentrations of P5 are shown in histogram form (fig. 4.6) and a table of P5 binding

Figure 4.3 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 517 bp pBR322 restriction fragment (*Eco*R I-*Rsa* I). The reactions were 100 μM in DNA base pairs (final concentration). In the footprinting lanes, P5 was allowed to equilibrate with the DNA one hour at 37°C. MPE·Fe(II) was then allowed to equilibrate with the DNA for 15 minutes. Cleavage was initiated by the addition of sodium acrobate to a final concentration of 1 mM and allowed to proceed for 15 minutes at 37°C.

Lanes 1-11 contain 5′-endlabeled DNA. Lanes 12-22 contain 3′-endlabeled DNA. Lanes 1 and 12 are intact DNA control lanes after incubation under reaction conditions. Lanes 2 and 13 are chemical sequencing lanes containing the products of Maxam-Gilbert G reactions. Lanes 3, 6, 9, 14, 17, and 20 contain 1:50 MPE·Fe(II)/bp. Lanes 4, 7, 10, 15, 18, and 21 contain 1:2 SE·Fe(II)/bp. Lanes 5, 8, 11, 16, 19, and 22 contain 1:2 EDTA·Fe(II)/bp. Lanes 6-8 and 17-19 contain 1:30 P5/bp. Lanes 9-11 and 20-22 contain 1:2 P5/bp.

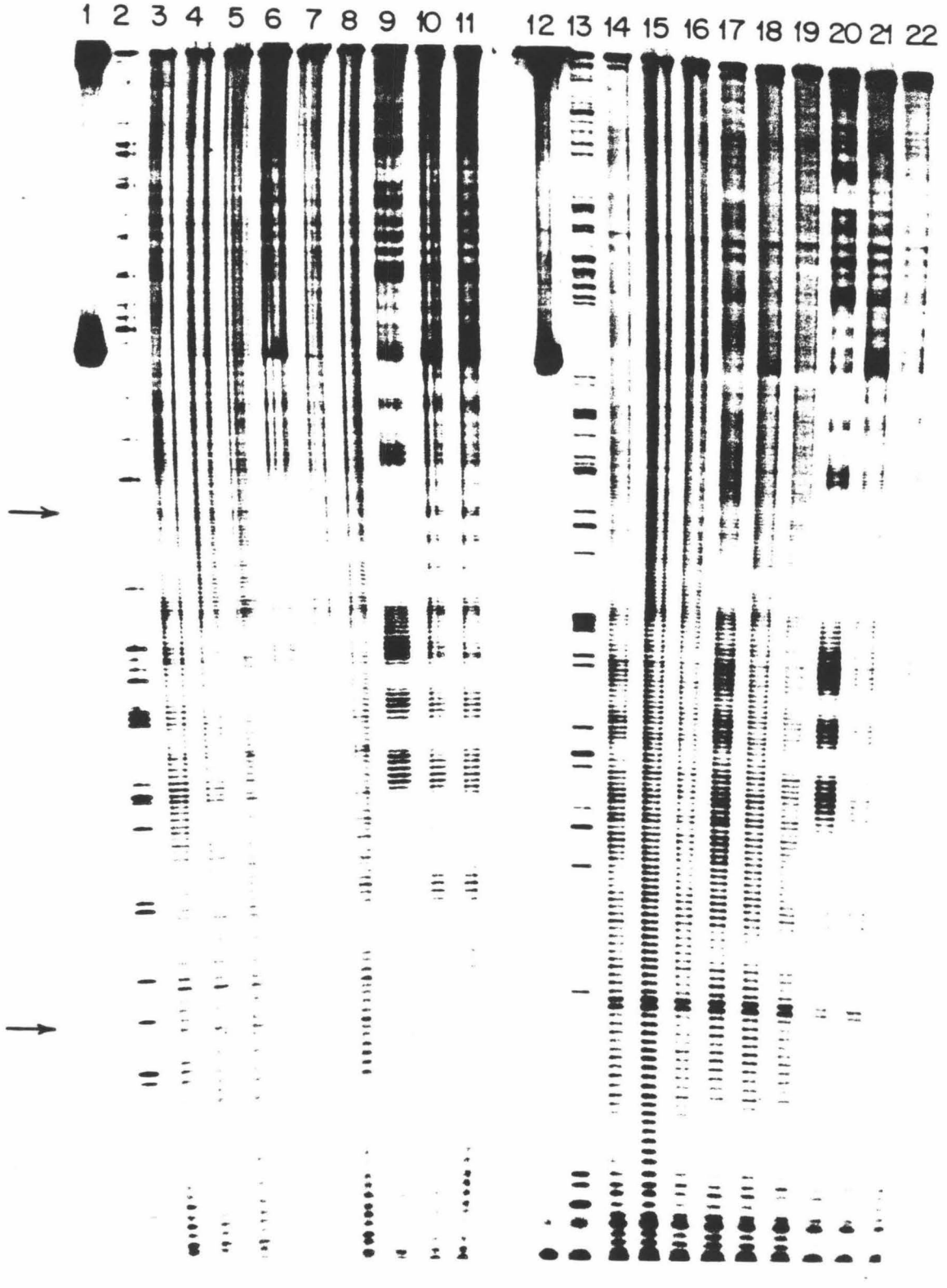
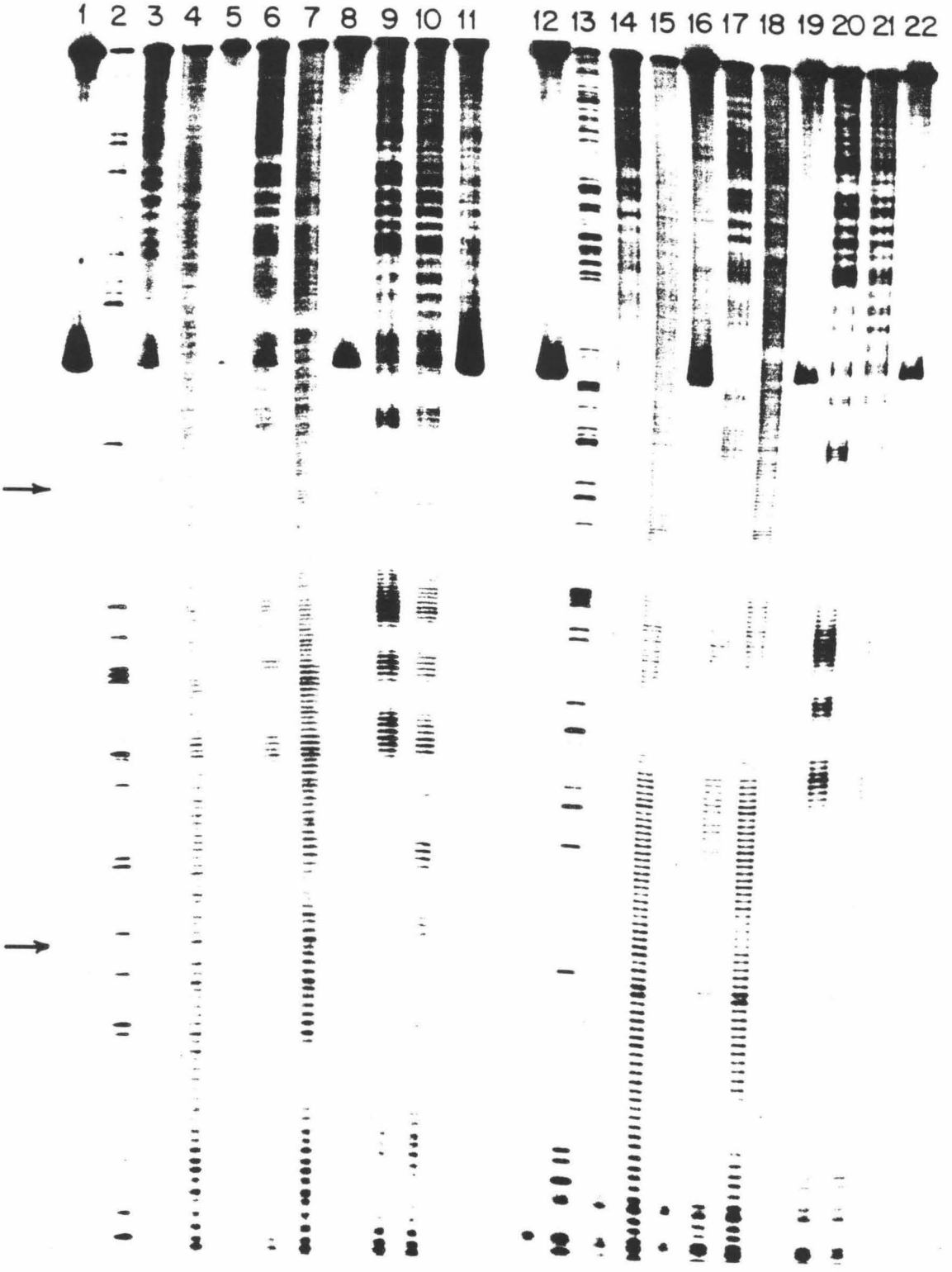

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

Figure 4.4 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 517 bp pBR322 restriction fragment (*Eco*R I-*Rsa* I). The reactions were 100 μM in DNA base pairs (final concentration). In the footprinting lanes, P5 was allowed to equilibrate with the DNA one hour at 37°C. MPE·Fe(II) was then allowed to equilibrate with the DNA for 15 minutes. Cleavage was initiated by the addition of dithiothreitol to a final concentration of 3 mM and allowed to proceed for 15 minutes at 37°C.

Lanes 1-11 contain 5′-endlabeled DNA. Lanes 12-22 contain 3′-endlabeled DNA. Lanes 1 and 12 are intact DNA control lanes after incubation under reaction conditions. Lanes 2 and 13 are chemical sequencing lanes containing the products of Maxam-Gilbert G reactions. Lanes 3, 6, 9, 14, 17, and 20 contain 1:50 MPE·Fe(II)/bp. Lanes 4, 7, 10, 15, 18, and 21 contain 1:2 SE·Fe(II)/bp. Lanes 5, 8, 11, 16, 19, and 22 contain 1:2 EDTA·Fe(II)/bp. Lanes 6-8 and 17-19 contain 1:30 P5/bp. Lanes 9-11 and 20-22 contain 1:2 P5/bp.


1 2 3 4 5 6 7 8 9 10 11
12 13 14 15 16 17 18 19 20 21 22

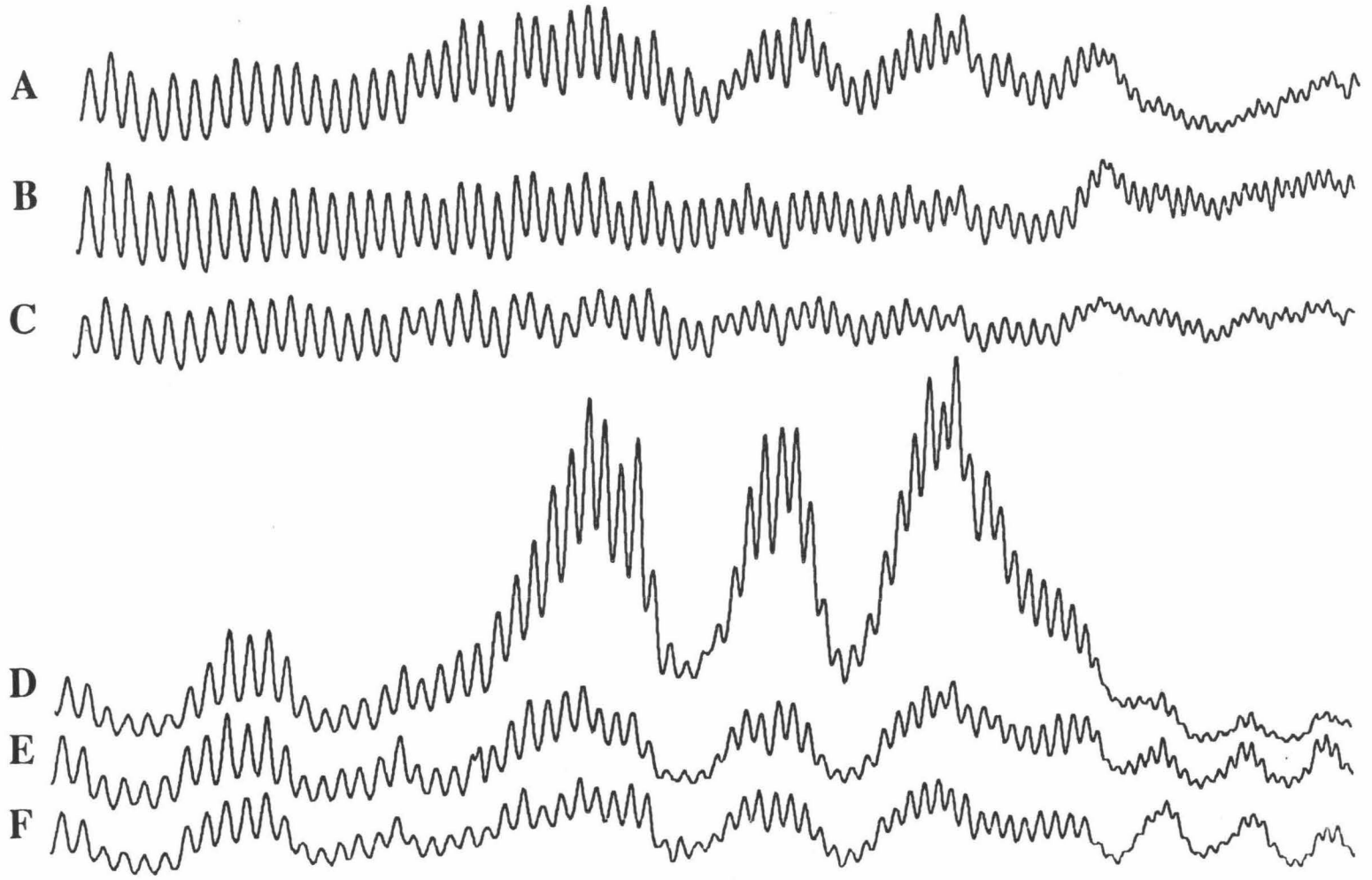


Figure 4.5 Densitometry plots of distance scanned versus absorbance for the autoradiogram shown in figure 4.3. (A) Trace of MPE·Fe(II) control (fig. 4.3 lane 14). (B) Trace of SE·Fe(II) control (fig. 4.3 lane 15). (C) Trace of EDTA·Fe(II) control (fig. 4.3 lane 16). (D) Trace of MPE·Fe(II) footprinting of P5 (1:2) (fig. 4.3 lane 20). (E) Trace of SE·Fe(II) footprinting of P5 (1:2) (fig. 4.3 lane 21). (F) Trace of EDTA·Fe(II) footprinting of P5 (1:2) (fig. 4.3 lane 22).

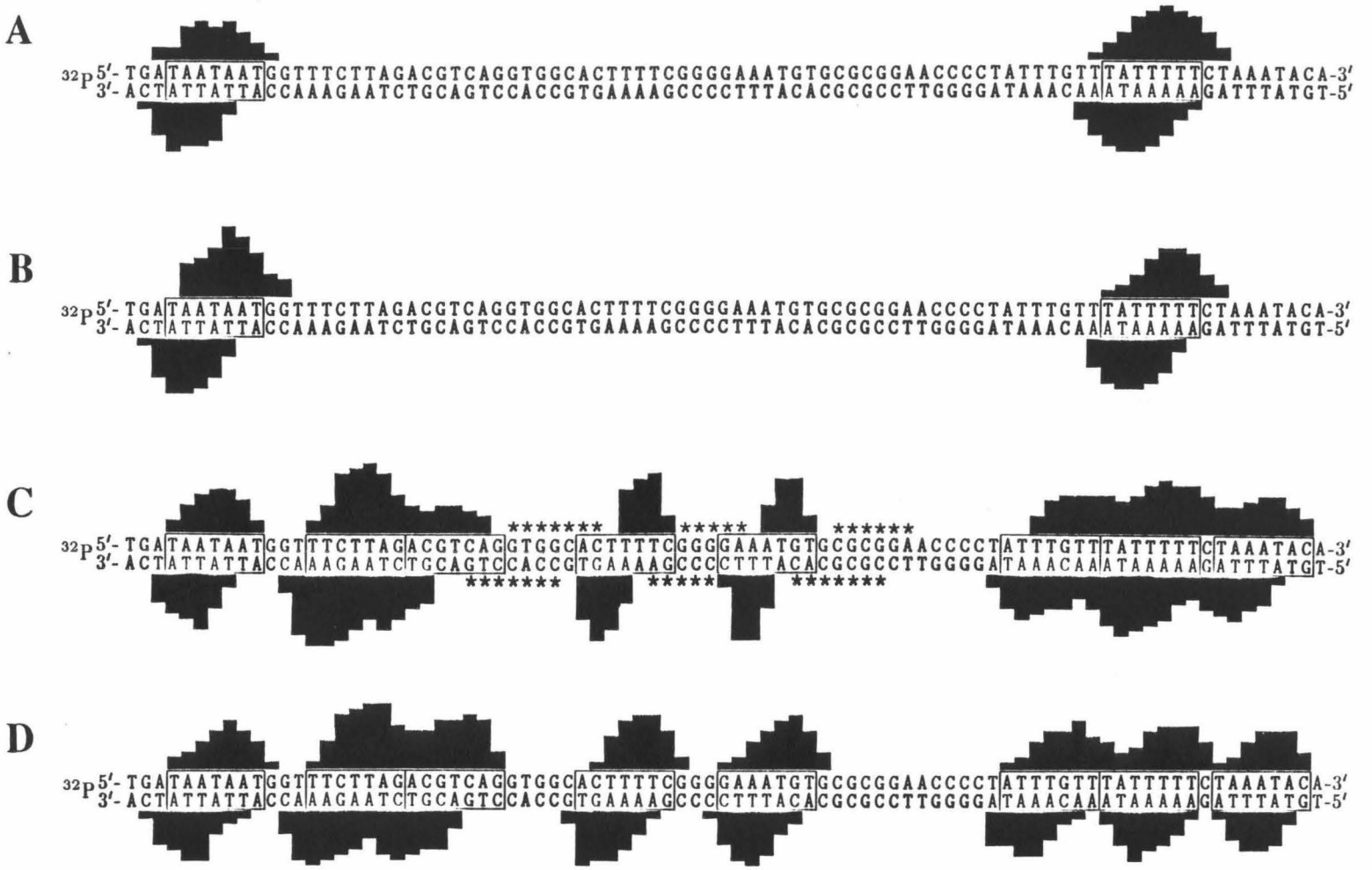
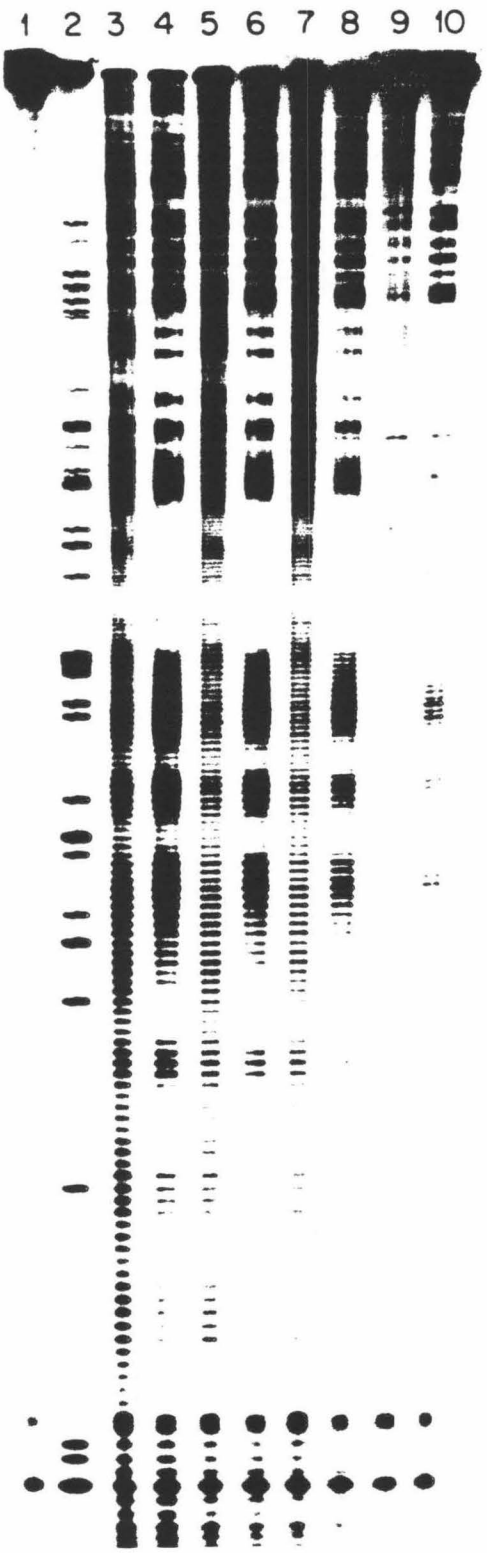


Figure 4.6 Histograms showing binding sites derived from P5 footprinting. Bars indicate cleavage inhibition and asterisks indicate cleavage enhancement. Boxes indicate binding site and size. (A) MPE·Fe(II) footprinting with 1:30 P5/bp. (B) EDTA·Fe(II) and SE·Fe(II) footprinting with 1:30 P5/bp. (C) MPE·Fe(II) footprinting with 1:2 P5/bp. (D) EDTA·Fe(II) and SE·Fe(II) footprinting with 1:2 P5/bp.

Figure 4.7 Autoradiogram of high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 517 bp pBR322 restriction fragment (*Eco*R I-*Rsa* I). The reactions were 100 μM in DNA base pairs (final concentration). In the footprinting lanes, P5 was allowed to equilibrate with the DNA one hour at 37°C. MPE·Fe(II) was then allowed to equilibrate with the DNA for 15 minutes at 37°C. The reactions were cooled to 0°C and cleavage was initiated by the addition of dithiothreitol to a final concentration of 3 mM.

Lanes 1-10 contain 3′-endlabeled DNA. Lane 1 is an intact DNA control lane after incubation under reaction conditions at 37°C. Lane 2 is a chemical sequencing lane containing the products of Maxam-Gilbert G reactions. Lanes 3-10 contain 1:50 MPE·Fe(II)/bp. Lanes 4, 6, 8, and 10 contain 1:2 P5/bp. Lanes 3 and 4 are cleavage at 4°C for 4 hours. Lanes 5 and 6 are cleavage at 20°C for 1 hour. Lanes 7 and 8 are cleavage at 37°C for 15 minutes. Lanes 9 and 10 are cleavage at 55°C for 5 minutes.


1 2 3 4 5 6 7 8 9 10

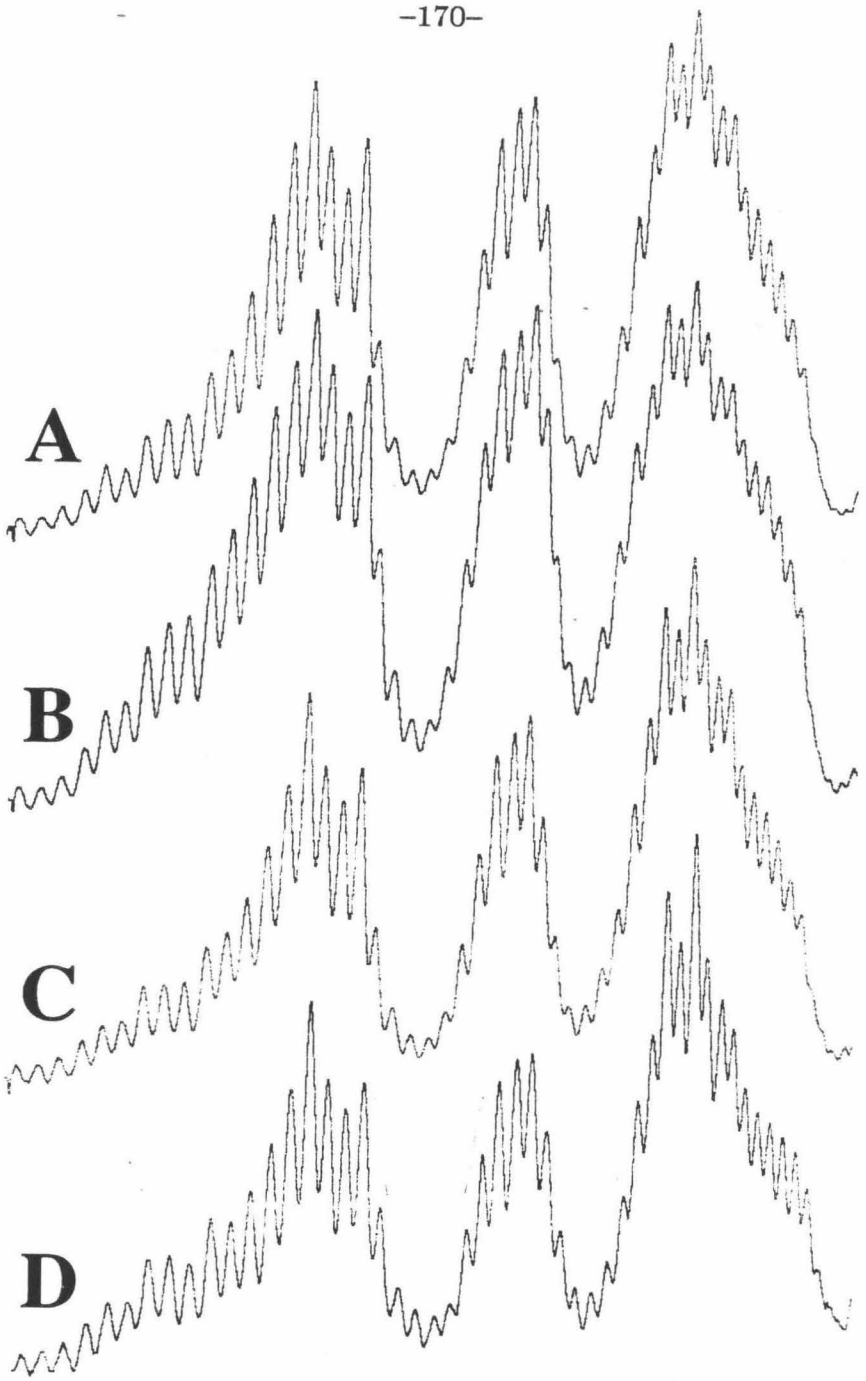


Figure 4.8 Densitometry plots of distance scanned versus absorbance for the autoradiogram shown in figure 4.7. (A) Trace of MPE·Fe(II) footprinting of P5 (1:2) at 4°C (fig. 4.7 lane 4). (B) Trace of MPE·Fe(II) footprinting of P5 (1:2) at 20°C (fig. 4.7 lane 6). (C) Trace of MPE·Fe(II) footprinting of P5 (1:2) at 37°C (fig. 4.7 lane 8). (D) Trace of MPE·Fe(II) footprinting of P5 (1:2) at 55°C (fig. 4.7 lane 10).

sites is given in table 4.1. With a P5:bp ratio of 1:2, extensive DNA cleavage is observed in three regions with MPE·Fe(II). Examination of the densitometry (fig. 4.5) shows no MPE·Fe(II) enhancement between the first two wells at the left, but obvious enhancement at three central sites. SE·Fe(II) and EDTA·Fe(II) detect all sites equivalently and give approximately the same amount of cleavage between each footprinting well. Varying the cleavage temperature from 4°C to 55°C did not eliminate the regions of enhanced cleavage seen with MPE·Fe(II) (fig. 4.7 and 4.8).

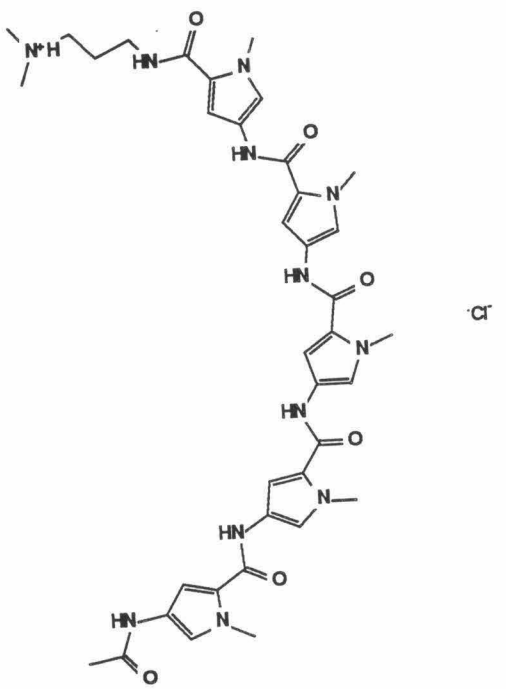


Figure 4.9 The structure of P5.

Cleavage Patterns. With EDTA·Fe(II) and 1 mM sodium ascorbate, very even DNA cleavage is observed. This is consistent with equal cleavage at each

Table 4.1: P5 Binding Sites the on 517 bp Fragment

P5/bp	Binding Sites
1:30	TAATAAT
	TATTTTT
1:2	TAATAAT
	TTCTTAG
	ACGTCAG
	ACTTTTC
	GAAATGT
	ATTTGTT
	TATTTTT
	TAAATAC

base and demonstrates that all of the sugars of DNA are equivalently accessible to the oxidative DNA cleaving species. Using a thiol reducing agent which can efficiently scavange radicals, EDTA·Fe(II) cleavage is inhibited. Cleavage inhibition is also seen with reactions containing acetate buffers or organic solvents. These results are consistent with the proposal that the active DNA cleaving species for the EDTA·Fe(II) containing compounds is a diffusible species such as hydroxyl radical which, in the case of EDTA·Fe(II), it is generated remotely and can be intercepted before reaching the DNA.

Because EDTA·Fe(II) generates its reactive species remotely, it is not interacting with the DNA and; therefore, it must be independent of the DNA sequence. In other words, the cleavage pattern EDTA·Fe(II) produces is the same as that produced by a molecule which binds DNA sequence neutrally. Consequently, deviations from the even cleavage pattern observed with EDTA·Fe(II) with another EDTA·Fe(II) DNA cleaving compound must result from an uneven

distribution of that cleaving molecule along the DNA.

With SE·Fe(II), a cleavage pattern equivalent to that of EDTA·Fe(II) with ascorbate is observed with either 3 mM DTT or 1 mM sodium ascorbate as the reductant. Because SE·Fe(II) is not inhibited by DTT, the cleaving species must be generated near the DNA verifying that SE·Fe(II) binds to the DNA. Because the even cleavage pattern is the same pattern observed by the non-DNA binding molecule EDTA·Fe(II), SE·Fe(II) interacts with the DNA sequence neutrally to single base pair resolution. This provides evidence that polyamines bind heterogeneous DNA sequence neutrally at the nucleotide level.

The uneven cleavage pattern seen with MPE·Fe(II) has been observed before.[41-43, 46] However, it is now possible to conclude that this pattern is due to variations in the binding constant of MPE·Fe(II) along the heterogeneous DNA polymer and is not a reflection of varying reactivity of the sugar backbone based on the even DNA cleavage pattern seen with both SE·Fe(II) and EDTA·Fe(II). The highest amounts of cleavage are seen in G·C regions and the lowest are found in A·T containing regions consistent with the known preference of intercalators for G·C base pairs.[111]

Footprinting. Footprinting of P5 at two concentrations showed cleavage inhibition at the same sites with each footprinting reagent. The sites observed are listed in table 4.1. They show that P5 binds to a 7 bp site as predicted from the n+1 rule.[112] This rule states that poly-*N*-methylpyrrolecarboxamides with n amides will bind n+1 base pairs. All the sites at low P5 binding density (1:30)

are exclusively A·T sites, but at higher binding density (1:2), sites containing G-C base pairs are observed. Polypyrrole compounds are known to prefer A·T rich DNA and their mode of DNA binding has been observed in a crystal structure of netropsin bound to DNA.[36-38] The fit of netropsin to DNA is this structure is very tight and it is not readily apparent how such an interaction could adjust to include a 2-amino group of guanine protruding into the minor groove. Because sites containing guanine are seen only at high concentrations, they must have a lower affinity and it may be that at the G·C containing sites P5 does not fit as deeply into the minor groove allowing room for the guanine 2-amino group.

Although each footprinting reagent shows the same P5 binding sites, there are some important differences between the different footprinting reagents. With sodium ascorbate as the reductant, SE·Fe(II) and EDTA·Fe(II) are equivalent. They both show very even control cleavage patterns and very well defined inhibition wells in the footprinting lanes. This allows the footprints to be easily interpreted. EDTA·Fe(II) has an advantage over other footprinting reagents because it does not bind to the DNA and consequently does not alter the structure of DNA. The disadvantage of generating radicals remotely from the DNA is that they have a much greater opportunity to react with substrates other than the DNA including the DNA ligand being footprinted. SE·Fe(II) on the other hand binds to DNA electrostatically and it does not appear to alter the DNA structure. Because SE·Fe(II) binds to DNA it is not noticeably inhibited by thiol reducing agents, organic buffers, or organic solvents such as methanol which are

sometimes used to overcome limited aqueous solubility of the small molecule being footprinted. By generating its DNA cleaving species near DNA, SE·Fe(II) is more likely to react with the DNA making it compatible with a wider range of reaction conditions than EDTA·Fe(II).

MPE·Fe(II) cleaves DNA at very low MPE·Fe(II):bp ratios and this minimizes the overall effect of intercalation on the DNA structure. However, at the sites that cleavage occurs, MPE·Fe(II) is either intercalated or partially intercalated.[67] This means that while MPE·Fe(II) does not change the global structure of the DNA, it has disrupted the local structure in the region in which it is reporting. MPE·Fe(II) also recognizes the 3° structure of DNA. This attribute is both an advantage and a disadvantage. It is a disadvantage in the respect that uneven cleavage in the control lane makes it difficult to assign footprints. If the molecule being footprinted disrupts the 3° structure of DNA, the altered enhancement/inhibition patterns compound the difficulty of footprint interpretation. On the other hand, 3° structure recognition is an advantage because it allows those regions of DNA which are being altered by the DNA ligand to be determined. An example of this effect is seen in the three regions of enhanced cleavage observed on the 517 bp fragment.

Cleavage Enhancement. A comparison of the gelscans (fig. 4.5) from the footprinting experiment on the 517 bp fragment shows greatly increased cleavage in three regions of DNA by MPE·Fe(II) with the addition of a high binding density of P5 (1:2). The ability of a bifunctional DNA binding/cleaving molecule such

as DNase I or MPE·Fe(II) to cut DNA is governed by two characteristics of the molecule; the binding affinity for a particular site and the efficiency of cleavage at that site. Without additional information, these two factors cannot be distinguished. Comparison of DNA cleavage with MPE·Fe(II) and EDTA·Fe(II) or SE·Fe(II) suggests that increased cleavage in the enhanced regions is due to an increased binding of MPE·Fe(II) and not simply an increase in the reactivity of those sugars to oxidative degradation.

The increased binding of MPE·Fe(II) can be explained two ways. P5 and MPE·Fe(II) both compete for the minor groove of DNA. One explanation for the enhanced cleavage is that the sites which show higher cleavage are the only sites where MPE·Fe(II) efficiently competes with P5 for the minor groove. However, the region between the first two P5 binding sites shown in figure 4.5 does not show cleavage enhancement and enhancement is not seen on other regions of the 517 bp fragment or with other DNA fragments at the same P5 concentration. This argues that enhanced cleavage is not merely a result of competition for the minor groove of DNA. A more likely possibility is that this region of enhanced cleavage results from a structural change induced by P5 which increases the binding affinity for MPE·Fe(II) to this region.

There are many examples in the literature of polypyrrole compounds binding to DNA and altering its structure. Kineoplast DNA has been shown to exist in a kinked form and the binding of distamycin A removes those kinks.[35] In another example, both distamycin A and netropsin have been shown to greatly

lengthen DNA isolated from certain sources, but not from others.[33,34] These studies also found a cooperative binding interaction between ethidium bromide and the poly-*N*-methylpyrrolecarboxamide compounds. Before we discuss the enhanced cleavage on this fragment, it should be noted that this region is just upstream of the –35 consensus region for the β-lactamase promoter[113] and includes a portion of the region indicated as the site of photo-cleavage of supercoiled pBR322 by Λ-tris(diphenylphenanthroline)cobalt(III) and attributed to Z DNA.[114] Because of its location, the structure of this region of DNA may be significant in the regulation of the β-lactamase gene.

The most obvious feature of this region is the inverted repeat of the primary structure (fig. 4.6). With an inverted repeat, the possibility of a cruciform structure arises. These structures have been mapped on supercoiled pBR322 but this particular site has not been seen.[115–118] To probe the possibility of cruciform formation or other modified DNA structure including single stranded and Z-form regions, the small molecule probes OsO_4 and diethyl pyrocarbonate and the enzyme S_1 nuclease were used;[115,116,119–121] however, none of the reagents showed any reactivity indicative of non-Watson-Crick DNA in the presence of P5 or P5 and ethidium. The reactivity of the DNA to dimethyl sulfate was also examined to probe possible structural alterations but revealed no unusual reactivity. Finally, MPE·Fe(II) footprinting of P5 on another DNA fragment containing an inverted repeat was carried out and no cleavage enhancement was observed. It is also important to remember that no cleavage enhancement or inhibition is

seen in these regions with SE·Fe(II) or EDTA·Fe(II), therefore the sugars must be equivalently accessible to the oxidative cleaving species in both the modified and unmodified forms. No indications of single stranded regions or any modified double stranded structure were detected by the small molecule or enzyme probes. The reactivity of the base sugars to cleavage mediated by SE·Fe(II) and EDTA·Fe(II) is as expected for double stranded Watson-Crick DNA. These results do not completely rule out radical DNA structures, but they suggest more modest DNA structural alterations such as local unwinding or conversion to another double stranded form with a higher affinity for MPE·Fe(II).

One structure consistent with these results is a junction between two forms of DNA. For instance, G·C regions of DNA are known to prefer the A-form and polypyrrole compounds are known to convert double stranded DNA to the B-form.[17–22] If P5 binding converts the A·T regions to B-form and the G·C regions remain A-form, the junction region between A and B form DNA may be less tightly stacked making intercalation more favorable.[122] It is also known that distamycin A removes bends from kineoplast DNA causing it to run slower in an acrylamide gel, but in the gel demonstrating this effect, distamycin A also causes another fragment to run faster (see ref. 35 fig. 5). Because distamycin causes only one fragment to run faster, this result is consistent with a sequence specific distamycin A induced bending of the DNA. If P5 bends the DNA adjacent to its binding site in the enhanced sequence, then the region of the bend may increase base pair separation in the minor groove of DNA allowing more exothermic intercalation by MPE·Fe(II).

With bent DNA, the degree of apparent bending has been found to be temperature dependent.[123, 124] To investigate the effect of temperature on the efficiency of enhanced cleavage, footprinting of P5 was carried out at four temperatures, 4°C, 20°C, 37°C, and 55°C (fig. 4.7). Comparison of the densitometry scans for these lanes shows that they are very similar although not identical for the entire temperature range. The important observation is that P5 demonstrates substantial enhanced cleavage between 4°C and 55°C (fig. 4.8). Therefore, P5 binding not only induces a structural change in an adjacent unbound region, but stabilizes that structural change over a 50°C temperature range.

This work has demonstrated that attachment of EDTA·Fe(II) to spermine creates a molecule, SE·Fe(II), which binds and cleaves DNA sequence neutrally. This lack of specificity provides evidence that at the nucleotide level polyamines recognize heterogeneous DNA independent of sequence. The lack of sequence specificity allows SE·Fe(II) to be used as a footprinting reagent. Comparison of SE·Fe(II) with two other small molecule footprinting reagents, EDTA·Fe(II) and MPE·Fe(II), found that both EDTA·Fe(II) and SE·Fe(II) produce an even cleavage pattern with well defined cleavage inhibition wells that facilitate binding site assignment. Because SE·Fe(II) is bound to the DNA and is not inhibited by thiols, organic solvents or buffer components in contrast to EDTA·Fe(II), it is compatible with a wider range of reaction conditions and, therefore, it is potentially a more useful footprinting reagent. MPE·Fe(II) produces a non-uniform cleavage pattern which results from an uneven binding affinity along the DNA

polymer. In the presence of a penta-*N*-methylpyrrolecarboxamide compound, MPE·Fe(II) shows regions of enhanced cleavage. These regions are most likely due to changes induced in the DNA structure by the binding of the penta-*N*-methylpyrrolecarboxamide molecule, possibly a change in either DNA form or DNA bending and were observed over a 50°C range.

Chapter 5

Experimental

MATERIALS AND METHODS

Nuclear magnetic resonance spectra were recorded on a a Varian Associates EM-390 (^{1}H=90.0), JEOL JNM-GX 400 (^{1}H=399.65) or Bruker WM (^{1}H=500.12) and are reported in parts per million downfield from tetramethylsilane. Infrared spectra were recorded on a Beckman 4210 or a Shimadzu IR-435 spectrophotometer. Ultraviolet-visible spectra were recorded on a Beckman Model 25 spectrophotometer, a Cary Model 219 spectrophotometer or a Hewlett-Packard Model 8451A spectrophotometer. Mass spectra were recorded using FAB techniques at the Midwest Center for Mass Spectrometry at the University of Nebraska. Agarose gels stained with ethidium bromide were photographed with Polaroid type 55 film and the negatives were scanned at 485 nm, with a Cary 219 spectrophotometer interfaced to an Apple II computer or on an LKB Ultroscan XL densitometer. Autoradiography of polyacrylamide gels or agarose gels with radiolabelled DNA was carried out at -70°C on Kodak X-Omat AR film.

Reagent grade chemicals were used without further purification unless otherwise stated. Chromatography was performed using EM Reagents Silica Gel (230-400 mesh). *N,N*-dimethylformamide (DMF) was dried over 4Å molecular seives. *N,N'*-carbonyldiimidazole was sublimed under reduced pressure prior to use. Ferrous ammonium sulfate was a Baker Analyzed Reagent. Dithiothreitol (DTT) was purchased from Boehring Mannheim. All non-aqueous reactions were run under argon with rigorous exclusion of water unless otherwise noted. All

water used in biological reactions and dilutions was deionized and distilled using a Corning Mega-Pure system. Aqueous (α^{32}P) dATP and dCTP, triethylammonium salts, 3000 Ci/mmol, were from Amersham and aqueous (γ^{32}P)ATP, 5000-9000 Ci/mmol, was from ICN. Nucleotide triphosphates were from Pharmacia. All enzymes were from Boehring Mannheim except the Klenow fragment of DNA polymerase I which was from New England Biolabs. Plasmid pBR322[125, 126] was grown in *Escherichia coli*, strain HB101, and isolated by standard procedures.[127] Calf thymus DNA from Sigma was sonicated, deproteinized, and extensively dialyzed.[71] λ-phage DNA was from Boehring Mannheim. Plasmid SV-CAT was generously provided by Nevis Fregien.

N-Methyl-4-nitropyrrole-2-carboxylic Acid 15. The monopyrrole nitro acid **15** was prepared by a modification of the procedure of Bialer *et al.*[90] A solution of 150 gm (1.37 mol) *N*-methylpyrrole-2-carboxylic acid in 900 mL acetic anhydride was stirred by means of an overhead stirrer and cooled to -25°C with a dry ice/acetone bath. A solution of 120 mL nitric acid at 0°C and 350 mL acetic anhydride at 0°C was prepared. This solution was added dropwise to the reaction mixture maintaining the reaction temperature at -25°C. The solution was stirred at -25°C for an additional hour and then poured onto 1300 mL of ice and water and stirred 1.5 h. After stirring, it was placed in a -20°C freezer for 12 h. The crystalline precipitate was collected and recrystallized from 1:1 acetic acid:water (v/v) to give 105 g (45%) monopyrrole nitro acid **15**. ^{1}H NMR (DMSO-d_6) δ 3.9 (s, 3H), 7.25 (d, J=2 Hz, 1H), 8.25 (d, J=2 Hz, 1H).

N-Methyl-4-nitropyrrole-2-carboxylic acid methyl ester 49. To solution of 16 g (0.09 mol) monopyrrole nitro acid **15** in 160 mL methanol, 16 mL sulfuric acid was added and the solution was refluxed 17 h. The solution was then cooled and the volume was reduced under vacuum. The white crystalline product was filtered off to give 15.5 g (90%) monopyrrole nitro methyl ester **49**.[90] ^{1}H NMR ($CDCl_3$) δ 3.85 (s, 3H), 4.0 (s, 3H), 7.35 (d, J=1.5 Hz, 1H), 7.55 (d, J=1.5 Hz, 1H).

N-Methyl-4-(N-methyl-4-nitropyrrole-2-carboxylamide)-pyrrole-2-carboxylic acid methyl ester 50. A solution of 15 g (0.08 mol) monopyrrole nitroester **49** in 130 ml DMF was reduced over 5% palladium on charcoal (1.5 g) at atmospheric pressure. The solution was filtered through Celite to remove the catalyst and 20 mL triethylamine (0.14 mol) was added. A mixture of 15 g *N*-methyl-4-nitropyrrole-2-carboxylic acid chloride (0.09 mol) was added in 20 mL DMF with stirring. (*N*-methyl-4-nitropyrrole-2-carboxylic acid chloride was prepared by refluxing the monopyrrole acid in thionyl chloride until clear (~1 h) and removing excess thionyl chloride under vacuum.) The mixture stirred 12 h, 700 mL ice and water was added and the product was filtered off. After washing with water twice and diethyl ether, the light-green product was dried yielding 24 g (95%) of the dipyrrole nitro ester **50**.[90] ^{1}H NMR (DMSO-d_6) δ 3.75 (s, 3H), 3.85 (s, 3H), 4.0 (s, 3H), 6.9 (d, J=1.5 Hz, 1H), 7.5 (d, J=1.5 Hz, 1H), 7.6 (d, J=2.0 Hz, 1H), 8.2 (d, J=1.5 Hz, 1H), 10.3 (s, 1H).

4-Nitro-tri-N-methylpyrrole-2-carboxylic acid methyl ester 14. A solution of 16.7 g (0.055 mol) dipyrrole nitro ester **50** in 500 mL DMF was

hydrogenated over 5% palladium on charcoal (1 g) at atmospheric pressure. The solution was filtered through Celite to remove the catalyst and 15 mL (0.11 mol) triethylamine was added. A solution of 11.1 g (0.065 mol) *N*-methyl-4-nitropyrrole-2-carboxylic acid chloride in 15 mL DMF was added. The reaction mixture stirred 12 h. A mixture of 600 mL ice and water was added and the product was filtered off. After washing with water twice and once with cold methanol, the light-green product was dried yielding 22 g (93%) of the tripyrrole nitro ester **14**.[90] ^{1}H NMR (DMSO-d_6) δ 3.75 (s, 3H), 3.83 (s, 3H), 3.85 (s, 3H), 4.0 (s, 3H), 6.9 (d, J=1.5 Hz, 1H), 7.05 (d, J=1.5 Hz, 1H), 7.25 (d, J=1.5 Hz, 1H), 7.5 (d, J=1.5 Hz, 1H), 7.6 (d, J=1.5 Hz, 1H), 8.2 (d, J=1.5 Hz, 1H), 10 (s, 1H), 10.35 (s, 1H).

4-Nitro-tri-N-methylpyrrole-2-carboxylic acid 21. To a suspension of 2 g (4.7 mMol) tripyrrole nitro ester **14** in 135 mL ethanol, a solution of 1.5 g (37.5 mMol) sodium hydroxide in 125 mL water was added and the mixture was refluxed for 3.5 h. The solution was cooled and concentrated to 150 mL under vacuum. After acidification with 6 N HCl, the product was filtered off, washed twice with water and once with cold methanol to give 1.9 g (95%) tripyrrole nitro acid **21**.[90] ^{1}H NMR (DMSO-d_6) δ 3.84 (s, 3H), 3.87 (s, 3H), 3.97 (s, 3H), 6.85 (d, J=1.5 Hz, 1H), 7.1 (d, J=1.5 Hz, 1H), 7.25 (d, J=1.5 Hz, 1H), 7.26 (d, J=1.5 Hz, 1H), 7.45 (d, J=1.5 Hz, 1H), 7.65 (d, J=1.5 Hz, 1H), 8.2 (d, J=1.5 Hz, 1H), 9.95 (s, 1H), 10.35 (s, 1H).

4-Nitro-tri-N-methylpyrrole-2-carboxamide dimethylpropylamine 53. A solution of 2.5 g (6.0 mMol) tripyrrole nitro acid **51**, 0.68 g (6.6 mMol)

3-dimethylaminopropylamine and 0.89 g (6.6 mMol) *N*-hydroxybenzotriazole in 10 mL DMF was cooled to 0°C, 1.36 g (6.6 mMol) dicyclohexylcarbodiimide was added and the solution was allowed to warm to 25°C and stir for 12 h. The DMF was removed under high vacuum, the residue was triturated twice with diethyl ether and then flash chromatographed on silica gel (concentrated aqueous ammonia:methanol, 3:97, v/v) to give 2.1 g (70%) tripyrrole nitro amine **53**.[70] IR (KBr) 3130, 2950, 1638, 1580, 1500, 1308, 1250 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.6 (m, 2H), 2.15 (s, 6H), 2.28 (t, J=6 Hz, 2H), 3.2 (m, 2H), 3.80 (s, 3H), 3.85 (s, 3H), 3.95 (s, 3H), 6.85 (d, J=1.5 Hz, 1H), 7.05 (d, J=1.5 Hz, 1H), 7.20 (d, J=1H), 7.27 (d, J=1.5 Hz, 1H), 7.6 (d, J=1.5 Hz, 1H), 8.05 (t, J=4 Hz, 1H), 8.15 (d, J=1.5 Hz, 1H), 9.95 (s, 1H), 10.35 (s, 1H).

EDTA triethylester 16. After suspending 10 g (0.034 mol) EDTA in dry ethanol, 1.5 mL sulfuric acid was added with stirring and the mixture was refluxed 24 h. A 50 mL saturated aqueous sodium bicarbonate solution was added and the reaction mixture was extracted with 250 mL dichloromethane. The organic layer was washed three times with saturated sodium bicarbonate, twice with water, dried over anhydrous sodium sulfate and concentrated to give 11 g (80%) crude EDTA-tetraethylester. The triethyl ester **16** was prepared according to the procedure of Hay and Nolan.[91] The 1.5 g (3.7 mMol) of the crude tetraester and 0.63 g (3.7 mol) cupric chloride dihydrate was dissolved in 100 mL water. A solution of 0.16 g (4.0 mMol) sodium hydroxide in 5 mL water was added with stirring at a rate to maintain the pH of the solution just

below 5. The solution was then bubbled with H_2S and filtered. The filtrate was concentrated under vacuum and purified by flash chromatography on silica gel (methanol:dichloromethane, 1:9, v/v) to give 0.89 g (64%) of the triethylester **16**.[70] ^{1}H NMR ($CDCl_3$) δ 1.3 (t, J=7 Hz, 9H), 2.74 (s, 4H), 3.3 (s, 2H), 3.4 (s, 2H), 3.5 (s, 4H), 4.1 (q, J=7 Hz, 6H).

EDTA triethylester-linker 17. To a solution of 5 g (0.013 mol) EDTA triethylester **16** and 1.52 g (0.13 mol) *N*-hydroxysuccinimide in 100 mL dioxane was added with stirring 2.7 g (0.013 mol) of dicyclohexylcarbodiimide in 20 mL dioxane. The solution was stirred for 12 h, filtered and the filtrate concentrated. The filtrate was then dissolved in 100 mL dimethoxyethane and added with stirring to a solution of 2 g (0.02 mol) 4-aminobutyric acid and 1.68 g (0.02 mol) sodium bicarbonate in 100 mL water. After stirring 12 h. the solution was concentrated under vacuum and purified by flash chromatography on silica gel (methanol:dichloromethane, 1:9, v/v) to give 4 g (65%) **17**.[70] ^{1}H NMR (DMSO-d_6) δ 1.19 (t, J=7 Hz, 9H), 1.63 (m, 2H), 2.2 (t, J=6 Hz, 2H), 2.7 (t, J=6 Hz, 2H), 3.1 (m, 2H), 3.19 (s, 2H), 3.45 (s, 2H), 3.53 (s, 4H), 4.08 (m, 6H), 8.0 (t, J=5 Hz, 1H).

Distamycin-EDTA triethylester 18. A solution of 200 mg (0.40 mMol) tripyrrole nitro amine **53** in 5 mL DMF was hydrogenated over 5% palladium on charcoal (60 mg) at atmospheric pressure. The solution was filtered throuth Celite to remove the catalyst. Activated EDTA-triethylester-linker **17** in 20 mL DMF was prepared by adding 80 mg *N,N'*-carbonyldiimidazole and allowing

the solution to stir 1 h. The reduced tripyrrole nitro amine was added and the solution stirred 12 h. The reaction mixture was concentrated under high vacuum, triturated three times with diethyl ether and purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 2:98, v/v) to give 300 mg (82%) **18**. ^{1}H NMR (DMSO-d_6) δ 1.19 (t, J=7 Hz, 9H), 1.60 (m, 2H), 1.75 (m, 2H), 2.13 (m, 2H), 2.2 (t, J=7 Hz, 2H), 2.26 (t, J=7 Hz, 2H), 2.7 (m, 4H), 3.1 (s, 2H), 3.2 (s, 2H), 3.45 (s, 2H), 3.55 (s, 4H), 3.84 (s, 3H), 3.88 (s, 3H), 3.90 (s, 3H), 4.08 (q, J=7 Hz, 6H), 6.80 (d, J=2 Hz, 1H), 6.86 (d, J=2 Hz, 1H), 7.16 (d, J=2 Hz, 1H), 7.18 (d, J=2 Hz, 1H), 7.22 (d, J=2 Hz, 1H), 8.00 (t, J=5 Hz, 1H), 8.05 (t, J=5 Hz, 1H), 9.80 (s, 1H), 9.88 (s, 1H), 10.37 (s, 1H).

Distamycin-EDTA 1. A solution of 39 mg (1.6 mMol) lithium hydroxide in 10 mL water was added with stirring to a solution of 70.5 mg (0.08 mMol) distamycin-EDTA triethylester **18** in 10 mL ethanol. The solution stirred 12 h and was concentrated under vacuum. The residue was purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 3:97, v/v) to give 55 mg (85%) distamycin-EDTA **1**.[70] ^{1}H NMR (DMSO-d_6) δ 1.73 (m, 2H), 1.85 (m, 2H), 2.3 (t, J=7 Hz, 2H), 2.72 (s, 6H), 3.05 (t, J=7 Hz, 2H), 3.13 (m, 2H), 3.22 (m, 2H), 3.40 (t, J=6 Hz, 2H), 3.45 (t, J=6 Hz, 2H), 3.80 (s, 3H), 3.82 (s, 3H), 3.83 (s, 3H), 3.93 (s, 4H), 4.04 (s, 2H), 4.17 (s, 2H), 6.91 (s, 1H), 6.93 (s, 1H), 7.05 (s, 1H), 7.18 (s, 1H), 7.22 (s, 1H), 7.26 (s, 1H), 8.25 (t, J=5 Hz, 1H), 8.86 (t, J=5 Hz, 1H), 9.95 (s, 1H), 9.97 (s, 1H), 10.15 (s, 1H); UV (H_2O) λ_{max} 238 nm, 305 nm.

4-Nitro-tetra-N-methylpyrrole-2-carboxylic acid methyl ester 37. A solution of 8 g (19 mMol) tripyrrole nitro ester **14** in 150 mL DMF was hydrogenated over 5% palladium on charcoal (1 g) at 52 psi hydrogen on a Parr rocker. The solution was filtered through Celite to remove the catalyst and 10 mL (136 mMol) triethylamine was added. A solution of 3.8 g (22 mMol) *N*-methyl-4-nitropyrrole-2-carboxylic acid chloride in 10 mL DMF was added and the reaction mixture stirred 12 h. After stirring, 600 mL ice and water was added and the product was filtered off. The product was washed twice with water, once with cold methanol and dried yielding 7.9 g (77%) of the tetrapyrrole nitro ester **37**. IR (KBr) 2950, 1700, 1660, 1580, 1440, 1400, 1305, 1245, 1105 cm^{-1}; ^1H NMR (DMSO-d_6) δ 3.79 (s, 3H), 3.84 (s, 3H), 3.86 (s, 3H), 3.87 (s, 3H), 3.97 (s, 3H), 6.91 (d, J=2 Hz, 1H), 7.06 (d, J=2 Hz, 1H), 7.08 (d, J=2 Hz, 1H), 7.25 (d, J=2 Hz, 1H), 7.28 (d, J=2 Hz, 1H), 7.47 (d, J=2 Hz, 1H), 7.60 (d, J=2 Hz, 1H), 8.18 (d, J=2 Hz, 1H), 9.93 (s, 1H), 9.99 (s, 1H), 10.28 (s, 1H).

4-Nitro-tetra-N-methylpyrrole-2-carboxylic acid 22. To a suspension of 730 mg (1.3 mMol) tetrapyrrole nitro ester **37** in 25 mL ethanol a solution of 0.5 g (12.5 mMol) sodium hydroxide in 25 mL water was added and the mixture was refluxed for 4 h. The solution was then cooled and concentrated to ~35 mL under vacuum. After acidification with 6 N HCl, the product was filtered off, washed three times with water and once with diethyl ether to give 630 mg (89%) tetrapyrrole nitro acid **22**. IR (KBr) 2950, 1645, 1580, 1540, 1430, 1400, 1305, 1245, 1110 cm^{-1}; ^1H NMR (DMSO-d_6) δ 3.83 (s, 3H), 3.86 (s, 3H), 3.88

(s, 3H), 3.97 (s, 3H), 6.87 (s, 1H), 7.06 (s, 1H), 7.07 (s, 1H), 7.25 (s, 1H), 7.28 (s, 1H), 7.43 (s, 1H), 7.60 (s, 1H), 8.80 (s, 1H), 9.90 (s, 1H), 9.99 (s, 1H), 10.29 (s, 1H); m/e 573 ($M+H^+$).

4-Nitro-tetra-N-methylpyrrole-2-carboxamide dimethylpropylamine 38. A solution of 500 mg (0.93 mMol) tetrapyrrole nitro acid **22**, 250 mg (1.9 mMol) *N*-hydroxybenzotriazole and 170 mg (1.7 mmol) 3-dimethylaminopropylamine in 5 mL DMF was cooled to 0°C and 230 mg (1.1 mMol) dicyclohexylcarbodiimide was added with stirring. After stirring 1 h at 0°C, the solution was allowed to warm to 25°C and stir for 12 h. The DMF was removed under high vacuum, the residue was triturated twice with diethyl ether and then purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 2:98, v/v) to give 520 mg (90%) tetrapyrrole nitro amine **38**. IR (KBr) 2950, 1630, 1580, 1530, 1465, 1430, 1308, 1255 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.60 (m, 2H), 2.20 (m, 2H), 2.30 (t, J=7 Hz, 2H), 3.2 (m, 2H), 3.80 (s, 3H), 3.85 (s, 3H), 3.88 (s, 3H), 3.95 (s, 3H), 6.85 (s, 1H), 7.07 (m, 2H), 7.20 (s, 1H), 7.27 (s, 1H), 7.31 (s, 1H), 8.05 (t, J=7 Hz, 1H), 8.1 (s, 1H), 9.92 (s, 1H), 9.97 (s, 1H), 10.03 (s, 1H), 10.35 (s, 1H); m/e 621 ($M+H^+$).

Tetra-N-methylpyrrolecarboxamide-EDTA triethylester 19. A solution of 240 mg (0.39 mMol) tetrapyrrole nitro amine **38** in 5 mL DMF was hydrogenated over 5% palladium on charcoal (60 mg) at atmopheric pressure. The solution was filtered through Celite to remove the catalyst. Activated EDTA triethylester-linker was prepared by dissolving 210 mg (0.46 mMol) EDTA

triethylester-linker **17** in 2 mL DMF, adding 80 mg (0.49 mMol) *N,N'*-carbonyldiimidazole and allowing the solution to stir 1 h. The reduced tetrapyrrole nitro amine was added to the activated EDTA triethylester-linker and the solution stirred 12 h. The reaction mixture was concentrated under high vacuum, triturated three times with diethyl ether and purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 3:97, v/v) to give 307 mg (77%) **19**. IR (KBr) 2943, 1738, 1649, 1582, 1539, 1533, 1465, 1434, 1403, 1256, 1204 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.18 (m, 9H), 1.62 (m, 2H), 1.72 (m, 2H), 2.15 (s, 6H), 2.25 (m, 4H), 2.70 (m, 2H), 2.73 (m, 2H), 3.14 (m, 2H), 3.19 (m, 4H), 3.46 (s, 2H), 3.51 (s, 4H), 3.80 (s, 3H), 3.84 (s, 3H), 3.85 (s, 3H), 3.86 (s, 3H), 4.06 (m, 6H), 6.83 (s, 1H), 6.88 (s, 1H), 7.05 (s, 1H), 7.16 (s, 1H), 7.18 (s, 1H), 7.24 (s, 2H), 8.00 (t, J=5 Hz, 1H), 8.06 (t, J=5 Hz, 1H), 9.80 (s, 1H), 9.88 (s, 1H), 9.90 (s, 1H), 9.93 (s, 1H); m/e 1034 ($M+H^+$).

Tetra-N-methylpyrrolecarboxamide EDTA 2. A solution of 47 mg (2.0 mMol) lithium hydroxide in 15 mL water was added with stirring to a solution of 101 mg (0.10 mMol) tetra-*N*-methylpyrrolecarboxamide-EDTA triethylester **19** in 10 mL ethanol. The solution stirred 12 h and was concentrated under vacuum. The residue was purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 3:97, v/v) to give 77 mg (83%) **2**. IR (KBr) 2950, 1620, 1580, 1540, 1460, 1435, 1400, 1260, 1200, 1100 cm^{-1}; ^1H NMR (DMSO-d_6+TFA) δ 1.75 (m, 2H), 1.86 (m, 2H), 2.28 (m, 2H), 2.80 (s, 6H), 3.09 (m, 2H), 3.16 (m, 4H), 3.27 (m, 4H), 3.73 (s, 4H), 3.85 (m, 14H), 4.04

(s, 2H), 6.89 (s, 1H), 6.95 (s, 1H), 7.08 (s, 1H), 7.16 (s, 1H), 7.18 (s, 1H), 7.23 (s, 2H), 8.17 (m, 1H), 8.43 (m, 1H), 9.85 (s, 1H), 9.91 (s, 1H), 9.94 (s, 1H); UV (H_2O) λ_{max} 238 nm, 310 nm; m/e 951 ($M+H^+$).

4-Nitro-penta-N-methylpyrrole-2-carboxylic acid methyl ester 54. A solution of 12.2 g (22 mMol) tetrapyrrole nitro ester **37** in 350 mL DMF was hydrogenated over 5% palladium on charcoal (2 g) at 52 psi hydrogen on a Parr rocker. The solution was filtered through Celite to remove the catalyst and 10 mL triethylamine was added. A solution of 4.7 g *N*-methyl-4-nitropyrrole-2-carboxylic acid chloride (27 mMol) in 10 mL DMF was added and the reaction mixture stirred 12 h. After stirring, 750 mL ice and water was added and the product was filtered off. The product was washed three times with water, once with cold methanol and dried yielding 13.5 g (90%) of the pentapyrrole nitro ester **54**. IR (KBr) 2950, 1770, 1660, 1590, 1520, 1440, 1390, 1305, 1250, 1100 cm^{-1}; ^1H NMR (DMSO-d_6) δ 3.78 (s, 3H), 3.85 (s, 9H), 4.02 (s, 3H), 6.91 (s, 1H), 7.10 (s, 3H), 7.45 (s, 1H), 7.60 (s, 1H), 8.14 (s, 1H), 9.86 (m, 3H), 10.20 (s, 1H); m/e 673 ($M+H^+$).

4-Nitro-penta-N-methylpyrrole-2-carboxylic acid 55. To a suspension of 5 g (7.4 mMol) pentapyrrole nitro ester **54** in 200 mL ethanol, a solution of 4 g (100 mMol) sodium hydroxide in 200 mL water was added and the mixture was refluxed for 4 h. The solution was cooled and reduced in volume to 250 mL under vacuum. After acidification with 6 N HCl, the product was filtered off, washed three times with water and once with diethyl ether to give 4.77 g (97%)

pentapyrrole nitro acid **55**. IR (KBr) 2950, 1650, 1590, 1520, 1435, 1395, 1305, 1250, 1105 cm^{-1}; ^1H NMR (DMSO-d_6) δ 3.85 (s, 12H), 4.00 (s, 3H), 6.88 (s, 1H), 7.10 (s, 3H), 7.30 (m, 3H), 7.45 (s, 1H), 7.61 (s, 1H), 8.20 (s, 1H), 9.90 (s, 1H), 9.95 (s, 1H), 10.00 (s, 1H), 10.30 (s, 1H).

4-Nitro-penta-N-methylpyrrole-2-carboxamide dimethylpropylamine 56. A solution of 1 g (1.5 mMol) pentapyrrole nitro acid **55**, 0.41 g (3.0 mMol) *N*-hydroxybenzotriazole and 0.19 g (1.9 mMol) 3-dimethylaminopropylamine in 175 mL DMF was cooled to 0°C and 0.47 g (2.3 mMol) dicyclohexylcarbodiimide was added with stirring. After stirring 1 h at 0°C, the solution was allowed to warm to 25°C and stir for 12 h. The DMF was removed under high vacuum, the residue was triturated three times with diethyl ether and then flash chromatographed on silica gel (concentrated aqueous ammonia:methanol, 3.5:96.5, v/v) to give 0.9 g (80%) pentapyrrole nitro amine **56**.[71] IR (KBr) 2945, 1627, 1580, 1530, 1465, 1435, 1307, 1260, 1110 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.65 (m, 2H), 2.20 (s, 6H), 2.30 (t, J=7 Hz, 2H), 3.2 (m, 2H), 3.83 (s, 3H), 3.90 (s, 3H), 3.91 (s, 3H), 3.92 (s, 3H), 4.00 (s, 3H), 6.85 (s, 1H), 7.10 (m, 3H), 7.20 (m, 3H), 7.27 (m, 2H), 7.31 (s, 1H), 8.10 (t, J=7 Hz, 1H), 8.20 (s, 1H), 9.92 (s, 1H), 9.97 (s, 1H), 10.03 (s, 1H); m/e 743 ($M+H^+$).

Penta-N-methylpyrrolecarboxamide-EDTA triethyl ester 20. A solution of 1 g (1.35 mMol) nitro amine **56** in 20 mL DMF was hydrogenated over 5% palladium on charcoal (200 mg) at atmospheric pressure. The mixture was filtered through Celite to remove the catalyst. Activated EDTA triethylester-linker was prepared by dissolving 620 mg (1.5 mMol) EDTA triethylester-linker

17 in 5 mL DMF and adding 240 mg (1.5 mMol) *N,N'*-carbonyldiimidazole and allowing the solution to stir 1 h. The reduced nitro compound was added and the solution stirred 12 h. The reaction mixture was concentrated under high vacuum, triturated three times with diethyl ether and purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 4:96, v/v) to give 0.6 g (40%) **20**.[71] IR (KBr) 2960, 1725, 1650, 1585, 1540, 1470, 1440, 1410, 1260, 1210, 1110 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.20 (t, J=7 Hz, 6H), 1.65 (m, 2H), 1.74 (m, 2H), 2.15 (s, 6H), 2.20 (m, 2H), 3.20 (m, 2H), 3.22 (s, 2H), 3.45 (s, 2H), 3.50 (s, 4H), 3.80 (s, 3H), 3.87 (m, 20H), 4.07 (m, 6H), 6.80 (s, 2H), 7.70 (m, 3H), 7.20 (s, 2H), 7.25 (s, 3H), 8.10 (t, J=7 Hz, 1H), 9.90 (s, 2H), 9.95 (s, 1H); m/e 1156 (M^+).

Penta-N-methylpyrrolecarboxamide-EDTA 3. A solution of 25 mg (1.0 mMol) lithium hydroxide in 8 mL water was added with stirring to a solution of 57.5 mg (0.05 mMol) penta-*N*-methylpyrrolecarboxamide-EDTA triethyl ester **20** in 8 mL ethanol. The solution was stirred 12 h and concentrated under vacuum. The residue was purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 3:97, v/v) to give 43 mg (80%) **3**.[71] IR (KBr) 2950, 1730, 1640, 1570, 1438, 1430, 1400, 1252, 1205 cm^{-1}; ^1H NMR (DMSO-d_6+TFA) δ 1.77 (m, 2H), 1.92 (m, 2H), 2.33 (t, J=7 Hz, 2H), 2.75 (s, 6H), 3.08 (m, 2H), 3.20 (m, 2H), 3.24 (m, 2H), 3.30 (m, 2H), 3.35 (m, 2H), 3.80 (s, 4H), 3.87 (m, 15H), 3.98 (s, 2H), 4.10 (s, 2H), 6.95 (m, 2H), 7.10 (m, 3H), 7.20 (s, 1H), 7.22 (s, 1H), 7.32 (s, 3H), 8.20 (s, 1H), 8.65 (s, 1H), 9.97 (m, 4H), 10.05 (s, 1H); UV (H_2O) λ_{max} 239 nm, 313 nm; m/e 1072 ($M+H^+$).

Tri-N-methylpyrrolecarboxylic acid-EDTA triethylester 23. A solution of 125 mg (0.30 mMol) tripyrrole nitro acid **21** in 2 mL DMF was hydrogenated at atmospheric pressure over 5% palladium on charcoal (10 mg). The mixture was filtered through Celite to remove the catalyst. A solution of 118 mg (0.26 mMol) EDTA triethylester-linker **17** and 45.5 mg (0.28 mMol) *N,N'*-carbonyldiimidazolein 4 mL DMF was allowed to stir 1 h. After 1 h, the reduced nitro acid was added and the solution stirred 12 h. The reaction was concentrated under vacuum and triturated three times with diethyl ether. The reaction mixture was loaded onto a flash silica gel column (methanol:dichloromethane, 1:9, v/v) and impurities were removed with a solvent gradient (methanol:dichloromethane, 1:9 to 1:3, v/v). The product was stripped from the column with methanol to give 114 mg (53%) of **23**. IR (KBr) 3100, 2940, 1730, 1650, 1570 (broad), 1435, 1400, 1200, 1020 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.19 (m, 9H), 1.71 (m, 2H), 2.25 (t, J=7 Hz, 2H), 2.70 (m, 2H), 2.72 (m, 2H), 3.22 (m, 2H), 3.19 (s, 2H), 3.45 (s, 2H), 3.50 (s, 4H), 3.85 (s, 9H), 4.06 (m, 6H), 6.43 (s, 1H), 6.68 (s, 1H), 6.98 (s, 1H), 7.03 (s, 1H), 7.18 (s, 1H), 7.23 (s, 1H), 8.03 (t, J=5 Hz, 1H), 9.67 (s, 1H), 9.90 (s, 3H); m/e 828 ($M+H^+$).

Hexa-N-methylpyrrolecarboxamide-EDTA triethyl ester 25. A solution of 21 mg (0.042 mMol) tripyrrole nitro amine **53** in 2 mL DMF was hydrogenated over 5% palladium on charcoal (10 mg) at atmospheric pressure. The solution was filtered through Celite to remove the catalyst. To the solution of reduced tripyrrole nitro amine 11.4 mg (0.084 mMol) *N*-hydroxybenzotriazole

and 35 mg (0.042 mMol) **23** was added with stirring. After the initial reactants dissolved, 10.5 mg (0.051 mMol) dicyclohexylcarbodiimide was added and the solution stirred 12 h. The DMF was removed under high vacuum, the residue was triturated with diethyl ether and then flash chromatographed (concentrated aqueous ammonia:(methanol:dichloromethane, 1:1, v/v), 1.5:98.5, v/v) to give 16 mg (30%) **25**. IR (KBr) 3100, 2950, 1734, 1653, 1576, 1558, 1540, 1436, 1256, 1204, 1026 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.18 (m, 9H), 1.62 (m, 2H), 1.75 (m, 2H), 2.14 (s, 6H), 2.23 (m, 4H), 2.70 (m, 2H), 2.73 (m, 2H), 3.14 (m, 2H), 3.17 (m, 2H), 3.19 (s, 2H), 3.46 (s, 2H), 3.51 (s, 4H), 3.80 (s, 3H), 3.84 (s, 3H), 3.85 (s, 3H), 3.87 (s, 9H), 4.06 (m, 6H), 6.83 (s, 1H), 6.88 (s, 1H), 7.06 (s, 1H), 7.09 (s, 3H), 7.16 (s, 1H), 7.18 (s, 1H), 7.25 (s, 3H), 8.00 (t, J=5 Hz, 1H), 8.06 (t, J=5 Hz, 1H), 9.80 (s, 1H), 9.88 (s, 1H), 9.89 (s, 1H), 9.94 (s, 2H); m/e 1278 ($M+H^+$).

Hexa-N-methylpyrrolecarboxamide-EDTA 4. A solution of 2.3 mg (0.10 mMol) lithium hydroxide in 0.48 mL water was added with stirring to a solution of 6 mg (0.005 mMol) hexa-*N*-methylpyrrolecarboxamide-EDTA triethyl ester **25** in 0.48 mL ethanol. The solution was stirred 12 h and was concentrated under vacuum. The residue was purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 2.5:97.5, v/v) to give 4.41 mg (78%) **4**. IR (KBr) 3020, 2930, 2850, 1730, 1640, 1585, 1465, 1435, 1405, 1260, 1205, 1100 cm^{-1}; ^1H NMR (DMSO-d_6+TFA) δ 1.75 (m, 2H), 1.85 (m, 2H), 2.29 (t, J=7 Hz, 2H), 2.80 (m, 6H), 3.09 (m, 2H), 3.17 (m, 4H), 3.26 (m, 2H), 3.31 (m,

2H), 3.74 (s, 4H), 3.85 (m, 18H), 3.91 (s, 2H), 4.07 (s, 2H), 6.89 (s, 1H), 6.96 (s, 1H), 7.09 (s, 2H), 7.11 (s, 2H), 7.16 (s, 1H), 7.17 (s, 1H), 7.23 (s, 4H), 8.16 (t, J=5 Hz, 1H), 8.41 (t, J=5 Hz, 1H), 9.82 (s, 1H), 9.90 (s, 2H), 9.94 (s, 3H); UV (H_2O) λ_{max} 240 nm, 314 nm; m/e 1195 ($M+H^+$).

Hepta-N-methylpyrrolecarboxamide-EDTA triethyl ester 26. A solution of 14.2 mg (0.023 mMol) tetrapyrrole nitro amine **38** in 2 mL DMF was hydrogenated over 5% palladium on charcoal (10 mg) at atmospheric pressure. The solution was filtered through Celite to remove the catalyst. To the solution of reduced tetrapyrrole nitro amine 6.2 mg (0.046 mMol) *N*-hydroxybenzotriazole and 19 mg (0.023 mMol) **23** was added with stirring. After the initial reactants dissolved, 5.7 mg (0.028 mMol) dicyclohexylcarbodiimide was added and the solution stirred 12 h. The DMF was removed under high vacuum, the residue was triturated with diethyl ether and then flash chromatographed (concentrated aqueous ammonia:(ethyl acetate:ethanol, 3:2, v/v), 4.4:95.6, v/v) to give 8.2 mg (27%) **26**. IR (KBr) 3020, 2920, 2850, 1735, 1645, 1580, 1540, 1480, 1435, 1405, 1260, 1200, 1105, 1020 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.18 (m, 9H), 1.62 (m, 2H), 1.71 (m, 2H), 2.18 (s, 6H), 2.23 (t, J=6 Hz, 2H), 2.30 (m, 2H), 2.70 (m, 4H), 3.12 (m, 2H), 3.18 (s, 2H), 3.46 (s, 2H), 3.80 (s, 3H), 3.83 (s, 3H), 3.86 (s, 15H), 4.06 (m, 6H), 6.83 (s, 1H), 6.88 (s, 1H), 7.06 (s, 1H), 7.07 (s, 1H), 7.09 (s, 3H), 7.16 (s, 1H), 7.18 (s, 1H), 7.24 (s, 5H), 8.00 (t, J=7 Hz, 1H), 8.09 (broad s, 1H), 9.81 (s, 1H), 9.91 (s, 1H), 9.95 (s, 4H); m/e 1400 (M^+).

Hepta-N-methylpyrrolecarboxamide-EDTA 5. A solution of 0.48 mg (0.021 mMol) lithium hydroxide in 1.5 mL water was added with stirring

to a solution of 2.8 mg (0.002 mMol) hepta-*N*-methylpyrrolecarboxamide-EDTA triethyl ester **26** in 2 mL ethanol. The solution was stirred 12 h and was concentrated under vacuum. The residue was purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 2.5:97.5, v/v) to give 1.98 mg (75%) **5**. IR (KBr) 3020, 2930, 2850, 1730, 1640, 1580, 1540, 1460, 1435, 1400, 1260, 1205, 1105 cm^{-1}; ^1H NMR (DMSO-d_6+TFA) δ 1.75 (m, 2H), 1.85 (m, 2H), 2.29 (t, J=7 Hz, 2H), 2.80 (m, 6H), 3.09 (m, 2H), 3.17 (m, 4H), 3.27 (m, 2H), 3.31 (m, 2H), 3.74 (s, 4H), 3.85 (m, 20H), 3.91 (s, 2H), 4.07 (s, 2H), 6.83 (s, 1H), 6.88 (s, 1H), 7.06 (s, 1H), 7.08 (s, 1H), 7.09 (s, 3H), 7.16 (s, 1H), 7.19 (s, 1H), 7.25 (s, 5H), 8.16 (t, J=6 Hz, 1H), 8.41 (t, J=6 Hz, 1H), 9.82 (s, 1H), 9.90 (s, 2H), 9.94 (s, 4H); UV (H_2O) λ_{max} 242 nm, 314 nm; m/e 1317 ($M+H^+$).

Tetra-N-methylpyrrolecarboxylic acid-EDTA triethylester 24. A solution of 58 mg (0.108 mMol) tetrapyrrole nitro acid **22** with 0.039 mL (0.28 mMol) triethylamine in 2 mL DMF was hydrogenated at atmospheric pressure over 5% palladium on charcoal (10 mg). The mixture was filtered through Celite to remove the catalyst. A solution of 55 mg (0.119 mMol) EDTA triethylester-linker **17** and 22 mg (0.136 mMol) *N,N'*-carbonyldiimidazole in 1 mL DMF was allowed to stir 1 h. After 1 h, the reduced nitro acid was added and the solution stirred 12 h. The reaction was concentrated under vacuum and triturated three times with diethyl ether. The reaction mixture was loaded onto a flash silica gel column (methanol:dichloromethane, 1:9, v/v) and impurities were removed with

a solvent gradient (methanol:dichloromethane, 1:9 to 1:3, v/v). The product was stripped from the column with methanol to give 53 mg (56%) of **24**. IR (KBr) 3100, 3010, 2930, 2850, 1730, 1650, 1580, 1545, 1435, 1400, 1235, 1200, 1100, 1060, 1020 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.17 (m, 9H), 1.72 (p, J=7 Hz, 2H), 2.24 (t, J=7 Hz, 2H), 2.70 (d, J=5 Hz, 2H), 2.73 (d, J=5 Hz, 2H), 3.13 (q, J=6 Hz, 2H), 3.19 (s, 2H), 3.46 (s, 2H), 3.51 (s, 4H), 3.83 (s, 9H), 3.86 (s, 3H), 4.06 (m, 6H), 6.42 (s, 1H), 6.88 (s, 1H), 6.99 (s, 1H), 7.05 (s, 2H), 7.17 (s, 1H), 7.23 (s, 1H), 7.25 (s, 1H), 8.02 (t, J=7 Hz, 1H), 9.68 (s, 1H), 9.85 (s, 1H), 9.92 (s, 2H); m/e 950 ($M+H^+$).

Octa-N-methylpyrrolecarboxamide-EDTA triethyl ester 27. A solution of 5.4 mg (0.0087 mMol) tetrapyrrole nitro amine **38** in 2 mL DMF was hydrogenated over 5% palladium on charcoal (1.5 mg) at atmospheric pressure. The solution was filtered through Celite to remove the catalyst. To the solution of reduced tetrapyrrole nitro amine 2.3 mg (0.017 mMol) *N*-hydroxybenzotriazole, 1.9 mg (0.0092 mMol) dicyclohexylcarbodiimide and 8 mg (0.0084 mMol) tetra-*N*-methylpyrrolecarboxylic acid-EDTA triethylester **24** was added with stirring. After the solution stirred 12 h, the DMF was removed under high vacuum, and the residue was triturated with diethyl ether and then flash chromatographed (concentrated aqueous ammonia:(methanol:dichloromethane, 1:1, v/v), 1:33, v/v) to give 4.4 mg (34%) **27**. IR (KBr) 3020, 2920, 2850, 1735, 1645, 1585, 1540, 1435, 1400, 1260, 1205, 1105 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.17 (m, 9H), 1.61 (p, J=7 Hz, 2H), 1.71 (p, J=7 Hz, 2H), 2.14 (s, 6H), 2.24 (t,

J=6 Hz, 4H), 2.68 (d, J=9 Hz, 2H), 2.71 (d, J=8 Hz, 2H), 3.14 (m, 2H), 3.19 (s, 2H), 3.51 (s, 2H), 3.60 (s, 4H), 3.80 (s, 3H), 3.84 (s, 3H), 3.87 (m, 18H), 4.07 (m, 6H), 6.83 (s, 1H), 6.89 (s, 1H), 7.09 (m, 14H), 7.16 (s, 1H), 7.19 (s, 1H), 7.25 (m, 14H), 8.01 (t, J=5 Hz, 1H), 8.07 (t, J=5 Hz, 1H), 9.81 (s, 1H), 9.89 (s, 1H), 9.90 (s, 1H), 9.96 (s, 5H); m/e 1523 ($M+H^+$).

Octa-N-methylpyrrolecarboxamide-EDTA 6. A solution of 1.0 mg (0.042 mMol) lithium hydroxide in 1 mL water was added with stirring to a solution of 3.4 mg (0.0022 mMol) octa-*N*-methylpyrrolecarboxamide-EDTA triethyl ester **27** in 1 mL ethanol. The solution was stirred 12 h and was concentrated under vacuum. The residue was purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 3:97, v/v) to give 2.60 mg (82%) **6**. IR (KBr) 3010, 2930, 2850, 1735, 1640, 1580, 1540, 1460, 1435, 1400, 1260, 1205, 1105 cm^{-1}; ^1H NMR (DMSO-d_6+TFA) δ 1.75 (p, J=6 Hz, 2H), 1.85 (p, J=7 Hz, 2H), 2.29 (t, J=7 Hz, 2H), 2.80 (d, J=5 Hz, 6H), 3.09 (m, 2H), 3.18 (m, 4H), 3.25 (m, 2H), 3.31 (m, 2H), 3.76 (s, 4H), 3.83 (s, 3H), 3.85 (s, 3H), 3.88 (m, 18H), 3.92 (s, 2H), 4.07 (s, 2H), 6.89 (s, 1H), 6.97 (s, 1H), 7.10 (s, 2H), 7.12 (s, 4H), 7.17 (s, 1H), 7.18 (s, 1H), 7.24 (s, 6H), 8.17 (t, J=6 Hz, 1H), 8.41 (t, J=6 Hz, 1H), 9.83 (s, 1H), 9.91 (s, 1H), 9.92 (s, 1H), 9.96 (s, 5H); UV (H_2O) λ_{max} 242 nm, 314 nm; m/e 1439 ($M+H^+$).

Nona-N-methylpyrrolecarboxamide-EDTA triethyl ester 28. A solution of 6.6 mg (0.0089 mMol) pentapyrrole nitro amine **56** in 2 mL DMF was hydrogenated over 5% palladium on charcoal (1 mg) at atmospheric pressure. The solution was filtered through Celite to remove the catalyst. To

the solution of reduced pentapyrrole nitro amine 2.3 mg (0.017 mMol) *N*-hydroxybenzotriazole, 1.9 mg (0.0092 mMol) dicyclohexylcarbodiimide and 8 mg (0.0084 mMol) tetra-*N*-methylpyrrolecarboxylic acid-EDTA triethylester **24** was added with stirring. After the solution stirred 12 h, the DMF was removed under high vacuum, and the residue was triturated with diethyl ether and then flash chromatographed (concentrated aqueous ammonia:(methanol:dichloromethane, 1:1, v/v), 1:33, v/v) to give 4.0 mg (29%) **28**. IR (KBr) 3010, 2920, 2850, 1730, 1645, 1580, 1540, 1435, 1400, 1260, 1200, 1105, 1020 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.21 (m, 9H), 1.66 (p, J=7 Hz, 2H), 1.76 (p, J=7 Hz, 2H), 2.18 (s, 6H), 2.29 (m, 4H), 2.72 (d, J=4 Hz, 2H), 2.75 (d, J=4 Hz, 2H), 3.21 (m, 6H), 3.21 (s, 6H), 3.45 (s, 2H), 3.52 (s, 4H), 3.83 (s, 3H), 3.87 (s, 3H), 3.90 (m, 21H), 4.08 (m, 6H), 6.84 (d, J=2 Hz, 1H), 6.91 (d, J=2 Hz, 1H), 7.09 (d, J=2 Hz, 1H), 7.11 (d, J=2 Hz, 1H), 7.13 (s, 5H), 7.23 (m, 7H), 8.06 (t, J=5 Hz, 1H), 8.22 (s, 1H), 9.78 (s, 1H), 9.86 (s, 1H), 9.88 (s, 1H), 9.93 (s, 6H).

Nona-N-methylpyrrolecarboxamide-EDTA 7. To 4 mg (0.0024 mMol) of nona-*N*-methylpyrrolecarboxamide-EDTA triethyl ester **28** and 0.5 mg (0.021 mMol) lithium hydroxide were added 2 mL of ethanol and 1 mL water. The solution was stirred 12 h and was concentrated under vacuum. The residue was purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 1:33, v/v) to give 3.5 mg (92%) **7**. IR (KBr) 3010, 2950, 2850, 1730, 1620, 1585, 1550, 1465, 1435, 1405, 1260, 1205, 1110 cm^{-1}; ^1H NMR (DMSO-d_6+TFA) δ 1.75 (p, J=7 Hz, 2H), 1.85 (p, J=7 Hz, 2H), 2.29 (t, J=7

Hz, 2H), 2.80 (d, J=4 Hz, 6H), 3.10 (m, 2H), 3.19 (m, 4H), 3.26 (m, 2H), 3.31 (m, 2H), 3.77 (s, 4H), 3.83 (s, 3H), 3.85 (s, 3H), 3.87 (s, 3H), 3.89 (m, 18H), 3.92 (s, 2H), 4.08 (s, 2H), 6.90 (s, 1H), 6.97 (s, 1H), 7.10 (s, 2H), 7.12 (s, 5H), 7.17 (s, 1H), 7.18 (s, 1H), 7.24 (s, 7H), 8.17 (t, J=6 Hz, 1H), 8.42 (t, J=6 Hz, 1H), 9.84 (s, 1H), 9.92 (s, 2H), 9.96 (s, 6H); UV (H_2O) λ_{max} 241 nm, 314 nm.

4-Nitro-di-N-methylpyrrole-2-carboxylic acid 51. To a suspension of 2 g (6.5 mMol) dipyrrole nitro ester **50** in 80 mL ethanol, a solution of 1.5 g (37.5 mMol) sodium hydroxide in 80 mL water was added and the mixture was refluxed for 3.5 h. The solution was cooled and concentrated to ~100 mL under vacuum. After acidification with 6 N HCl, the product was filtered off and washed four times with water and dried to give 1.99 g (104%) dipyrrole nitro acid **51.** ^{1}H NMR (DMSO-d_6) δ 3.83 (s, 3H), 3.94 (s, 3H), 6.83 (d, J=2 Hz, 1H), 7.41 (s, 1H), 7.56 (d, J=2 Hz, 1H), 8.16 (d, J=2 Hz, 1H), 10.24 (s, 1H).

4-Nitro-di-N-methylpyrrole-2-carboxamide dimethylpropylamine 52. To a solution of 78 mg (0.27 mMol) dipyrrole nitro acid **21**, 29.9 mg (0.29 mMol) 3-dimethylaminopropylamine and 71.9 mg (0.53 mMol) *N*-hydroxybenzotriazole in 2 mL DMF was added 60.4 mg (0.29 mMol) dicyclohexylcarbodiimide and the solution was allowed to stir for 12 h. The DMF was removed under high vacuum, the residue was triturated twice with diethyl ether and then flash chromatographed on silica gel (concentrated aqueous ammonia:methanol, 3:97, v/v) to give 84 mg (84%) dipyrrole nitro amine **52.** IR (KBr) 3130, 2950, 1638, 1580, 1500, 1308, 1250 cm^{-1}; ^{1}H NMR (DMSO-d_6) δ

1.62 (m, 2H), 2.15 (s, 6H), 2.26 (t, J=7 Hz, 2H), 3.20 (t, J=7 Hz, 2H), 3.81 (s, 3H), 3.96 (s, 3H), 6.82 (d, J=2 Hz, 1H), 7.20 (d, J=2H), 7.57 (d, J=2 Hz, 1H), 8.11 (t, J=5 Hz, 1H), 8.16 (d, J=2 Hz, 1H), 10.35 (s, 1H); m/e 377 ($M+H^+$).

4-Acetamide-di-N-methylpyrrole-2-carboxamide dimethylpropylamine 8. A solution of 200 mg (0.51 mMol) nitro di-*N*-methylpyrrolecarboxamide dimethylpropylamine **52** in 5 ml DMF was hydrogenated at atmospheric pressure over 5% palladium on charcoal (40 mg). The catalyst was filtered off and 300 μL (4.08 mMol) triethylamine followed by 100 μL (1.15 mMol) acetyl chloride was added. The reaction mixture was stirred overnight and then concentrated under high vacuum. The residue was triturated with diethyl ether and flash chromatographed (concentrated aqueous ammonia:methanol, 2:98, v/v) to give 175 mg (88%) **8**. IR (KBr) 3100, 2940, 1640, 1580, 1530, 1435, 1400, 1260, 1205, 1095 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.61 (p, J=7 Hz, 2H), 1.97 (s, 3H), 2.13 (s, 6H), 2.23 (t, J=7 Hz, 2H), 3.17 (t, J=6 Hz, 2H), 3.80 (s, 3H), 3.82 (s, 3H), 6.81 (s, 1H), 6.83 (s, 1H), 7.14 (s, 1H), 7.18 (s, 1H), 8.06 (t, J=5 Hz, 1H), 9.81 (s, 1H), 9.84 (s, 1H); UV (H_2O) λ_{max} 235 nm (21,600), 296 nm (22,800); m/e 389 ($M+H^+$).

4-Acetamide-tri-N-methylpyrrole-2-carboxamide dimethylpropylamine 9. A solution of 300 mg (0.60 mMol) tri-*N*-methylpyrrolecarboxamide dimethylpropylamine **53** in 2 mL DMF at atmospheric pressure over 5% palladium on charcoal (50 mg), the catalyst was filtered off and acetic acid imidazolide (prepared by adding 300 mg (1.85 mMol) carbonyldiimidazole to 105 μL (1.8

mMol) acetic acid in 2 mL DMF and stirring 2H) was added. The reaction mixture was stirred overnight and concentrated under high vacuum. The residue was triturated with diethyl ether and after flash chromatography (concentrated aqueous ammonia:methanol, 2:98, v/v) gave 230 mg (75%) **9**. IR (KBr) 3100, 2930, 1640, 1580, 1530, 1435, 1400, 1260, 1200, 1100 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.87 (t, J=7.0 Hz, 2H), 2.78 (s, 3H), 3.06 (t, J=7 Hz, 2H), 3.25 (q, J=6.0 Hz, 2H), 3.82 (s, 3H), 3.83 (s, 3H), 3.85 (s, 3H), 6.88 (s, 1H), 6.94 (s, 1H), 7.06 (s, 1H), 7.15 (s, 1H), 7.19 (s, 1H), 7.24 (s, 1H), 8.18 (t, J=5 Hz, 1H), 9.86 (s, 1H), 9.92 (s, 2H); UV (H_2O) λ_{max} 237 nm (28,000), 303 nm (33,000); m/e 511 ($M+H^+$).

4-Acetamide-tetra-N-methylpyrrole-2-carboxamide dimethylpropylamine 10. A solution of 100 mg (0.16 mMol) tetra-*N*-methylpyrrolecarboxamide dimethylpropylamine **38** in 5 mL DMF at atmospheric pressure over 5% palladium on charcoal (20 mg). The catalyst was filtered off and 134 μL (1.82 mMol) triethylamine followed by 40 μL (0.46 mMol) acetyl chloride was added. The reaction mixture was stirred overnight and then concentrated under high vacuum. The residue was triturated with diethyl ether and flash chromatographed (concentrated aqueous ammonia:methanol, 2.6:97.4, v/v) to give 67 mg (66%) **10**. IR (KBr) 2910, 1730, 1640, 1580, 1545, 1535, 1435, 1400, 1255, 1200, 1100 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.86 (p, J=7.7 Hz, 2H), 2.77 (s, 6H), 3.08 (t, J=1.7 2H), 3.25 (q, J=6.0 Hz, 2H), 3.82 (s, 3H), 3.84 (s, 3H), 3.85 (s, 3H), 3.86 (s, 3H), 6.88 (s, 1H), 6.95 (s, 1H), 7.08 (s, 2H), 7.15 (s, 1H), 7.19 (s,

1H), 7.24 (s, 2H), 8.18 (t, J=7 Hz, 1H), 9.92 (s, 1H), 9.93 (s, 1H), 9.95 (s, 1H); UV (H_2O) λ_{max} 238 nm (30,000), 308 nm (40,000); m/e 633 ($M+H^+$).

4-Acetamide-penta-N-methylpyrrole-2-carboxamide dimethylpropylamine 11. A solution of 155mg (0.21 mMol) penta-*N*-methylpyrrolecarboxamide dimethylpropylamine **56** in 30 mL DMF at atmospheric pressure over 5% palladium on charcoal (36 mg). The catalyst was filtered off and 300 μL (4.08 mMol) triethylamine followed by 40 μL (0.46 mMol) acetyl chloride was added. The reaction mixture was stirred overnight and then concentrated under high vacuum. The residue was triturated with diethyl ether and flash chromatographed (concentrated aqueous ammonia:methanol, 2.5:97.5, v/v) to give 95 mg (62%) **11**. IR (KBr) 3100, 3010, 2920, 2850, 1730, 1640, 1580, 1545, 1435, 1400, 1255, 1200, 1100 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.86 (m, 2H), 1.98 (s, 2H), 2.77 (s, 6H), 3.06 (t, J=7 Hz, 2H),3.25 (m, 2H), 3.78 (s, 3H), 3.82 (s, 3H), 3.84 (s, 3H), 3.89 (s, 3H), 6.88 (s, 1H), 6.95 (s, 1H), 7.08 (s, 2H), 7.09 (s, 1H), 7.15 (s, H), 7.19 (s, H), 7.24 (s, 3H), 8.18 (t, J=7 Hz, 1H), 9.86 (s, 1H), 9.93 (s, 2H), 9.96 (s, 2H); UV (H_2O) λ_{max} 238 nm (35,000), 310 nm (46,000); m/e 755 ($M+H^+$).

4-Acetamide-tri-N-methylpyrrole-2-carboxylic acid methyl ester 29. A solution of 100 mg (0.234 mMol) 4-nitro-tri-*N*-methylpyrrolecarboxylic acid methyl ester **14** in 4 mL DMF at atmospheric pressure over 5% palladium on charcoal (20 mg) at 37°C for 3 h. The catalyst was filtered off through a Celite plug. To the stirring reaction mixture 130 μL (94 mg, 4 eq) triethylamine followed by 40 μL (44 mg, 2 eq) acetyl chloride was added. After removing

the solvent under high vacuum and trituration with diethyl ether, the reaction mixture was purified by flash chromatography on silica gel to provide 60 mg (0.11 mMol, 99% based on recovered starting material) **29**. IR (KBr) 3100, 2950, 1710, 1650, 1580, 1435, 1405, 1250, 1200, 1100 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.97 (s, 3H), 3.83 (m, 9H), 6.83 (m, 2H), 7.04 (s, 1H), 7.15 (s, 1H), 7.23 (s, 1H), 7.24 (s, 1H), 9.80 (s, 1H), 9.89 (s, 2H).

4-Acetamide-tri-N-methylpyrrole-2-carboxylic acid 30. A solution of 200 mg (0.45 mMol) 4-acetamide-tri-*N*-methylpyrrolecarboxylic acid methyl ester **29** and 55 mg (1.38 mMol) sodium hydroxide in 5 mL water and 5 mL ethanol was refluxed for 4 h. The ethanol was removed under reduced pressure. The solution was cooled in an ice bath and 0.75 mL (1.50 mMol) 2 N HCl was added. The precipitate was filtered off and washed once with water, twice with ether and dried to afford 167 mg (87%) **30**. IR (KBr) 3100, 2930, 2850, 1640, 1580, 1435, 1400, 1235, 1205, 1110 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.97 (s, 3H), 3.82 (s, 3H), 3.83 (s, 3H), 3.84 (s, 3H), 6.85 (s, 2H), 7.05 (s, 1H), 7.16 (s, 1H), 7.24 (s, 1H), 7.44 (s, 1H), 9.84 (s, 1H), 9.91 (s, 2H); m/e 427 ($M+H^+$).

4-Acetamide-hexa-N-methylpyrrole-2-carboxamide dimethylpropylamine 12. A solution of 35 mg (0.70 mMol) tri-*N*-methylpyrrolecarboxamide dimethylpropylamine **53** in 5 mL DMF at atmospheric pressure over 5% palladium on charcoal (10 mg). The catalyst was filtered off through a Celite plug and 30 mg (0.07 mMol) 4-acetamide-tri-*N*-methylpyrrole-2-carboxylic acid **30**, 16 mg (0.078 mMol) dicyclohexylcarbodiimide and 19 mg (0.141 mMol) *N*-hydroxybenzotriazole were added. The reaction mixture was stirred overnight

and then concentrated under high vacuum. The residue was triturated with diethyl ether and flash chromatographed (concentrated aqueous ammonia:(ethyl acetate:ethanol, 3:2, v/v), 6:94, v/v) to give 31 mg (50%) **12**. IR (KBr) 3010, 2920, 2850, 1730, 1640, 1580, 1540, 1435, 1400, 1255, 1200, 1100 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.83 (m, 2H), 1.98 (s, 3H), 3.00 (m, 2H), 3.26 (m, 2H), 3.81 (s, 3H), 3.84 (s, 3H), 3.86 (s, 3H), 3.87 (s, 9H), 6.88 (s, 1H), 6.94 (s, 1H), 7.08 (s, 2H), 7.10 (s, 2H), 7.15 (s, 1H), 7.19 (s, 1H), 7.25 (s, 4H), 8.17 (t, J=5 Hz, 1H), 9.85 (s, 1H), 9.93 (s, 2H), 9.96(s, 2H); UV (H_2O) λ_{max} 240 nm (38,000), 314 nm (53,000); m/e 877 ($M+H^+$).

4-Acetamide-hepta-N-methylpyrrole-2-carboxamide dimethylpropylamine 13. A solution of 60 mg (0.097 mMol) tetra-*N*-methylpyrrolecarboxamide dimethylpropylamine **38** in 2 mL DMF at atmospheric pressure over 5% palladium on charcoal (15 mg). The catalyst was filtered off through a Celite plug and 46.9 mg (0.109 mMol) 4-acetamide-tri-*N*-methylpyrrole-2-carboxylic acid **30**, 25 mg (0.121 mMol) dicyclohexylcarbodiimide and 33 mg (0.244 mMol) *N*-hydroxybenzotriazole were added. The reaction mixture was stirred overnight and then concentrated under high vacuum. The residue was triturated with diethyl ether and flash chromatographed in (concentrated aqueous ammonia:(ethyl acetate:ethanol, 3:2, v/v), 7:93, v/v) to give 66.2 mg (68%) **13**. IR (KBr) 3020, 2910, 2850, 1735, 1645, 1580, 1545, 1435, 1400, 1260, 1200, 1100 cm^{-1}; ^1H NMR (DMSO-d_6) δ 1.85 (m, 2H), 1.98 (s, 2H), 2.81 (s, 6H), 3.07 (t, J=2 Hz, 2H), 3.25 (t, J=6 Hz, 2H), 3.82 (s, 3H), 3.84 (s, 3H), 3.86 (s, 3H), 3.87 (s, 12H), 6.88

(s, 1H), 6.96 (s, 1H), 7.08 (s, 2H), 7.10 (s, 3H), 7.15 (s, 1H), 7.19 (s, 1H), 7.26 (s, 5H), 8.18 (t, J=7 Hz, 1H), 9.84 (s, 1H), 9.92 (s, 2H), 9.94 (s, 4H); m/e 999 ($M+H^+$).

Tri-N-methylpyrrolecarboxamide-4-NO_2-*t*-boc amine 33. To a solution of 1 g (1.85 mMol) tripyrrole nitro amine **32**[71] in 4 mL DMF was added 0.48 gm (2.2 mMol) di-*tert*-butyl-dicarbonate. The reaction was allowed to proceed 3 h and then was concentrated under high vacuum. The residue was triturated three times with anhydrous diethyl ether and flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 1:500, v/v) to afford 0.87g (75%) **33**. ^1H NMR (DMSO-d_6) δ 1.36 (s, 9H), 1.50 (m, 2H), 1.60 (m, 2H), 2.12 (s, 3H), 2.27 (m, 4H), 2.93 (m, 2H), 3.17 (m, 2H), 3.77 (s, 3H), 3.85 (s, 3H), 3.98 (s, 3H), 6.79 (t, J=5 Hz, 1H), 6.82 (s, 1H), 7.02 (s, 1H), 7.18 (s, 1H), 7.27 (s, 1H), 7.59 (s, 1H), 8.03 (t, J=5 Hz, 1H), 8.19 (s, 1H), 9.93 (s, 1H), 10.28 (s, 1H); m/e 642 ($M+H^+$).

Bisdistamycin *t*-boc amine fumaramide 34. A solution of 150 mg (0.23 mMol) **33** in 1 mL DMF was hydrogenated at atmospheric pressure over 5% palladium on charcoal (30 mg). After hydrogenation, the catalyst was filtered off through a plug of Celite, 13.6 mg (0.117 mMol) fumaric acid, 57 mg (0.276 mMol) dicyclohexylcarbodiimide and 50 mg (0.37 mMol) *N*-hydroxybenzotriazole were added and the reaction was allowed to continue for 24 h. The solvent was removed under vacuum and the residue was triturated with diethyl ether. Chromatography was attempted (concentrated aqueous ammonia:methanol, 6.5:93.5,

v/v) but produced mixed fractions. The solids were taken up in solvent (concentrated aqueous ammonia:methanol, 1:250, v/v) and 45 mg pure product precipitated on standing. The supernatant was purified by flash chromatography on silica gel ((concentrated aqueous ammonia:methanol, 1:250, v/v) then (concentrated aqueous ammonia:(ethyl acetate:ethanol, 3:2, v/v), 1:50, v/v)) to yield 32 mg **34**. The total yield from the reaction was 77mg (50%) **34**. ^{1}H NMR (DMSO-d_6) δ 1.36 (s, 9H), 1.52 (m, 2H), 1.62 (m, 2H), 2.14 (s, 3H), 2.30 (m, 4H) 2.93 (m, 2H), 3.18 (q, J=6 Hz, 2H), 3.80 (s, 3H), 3.85 (s, 3H), 3.88 (s, 3H), 6.80 (broad s, 1H), 6.85 (s, 1H), 6.96 (s, 1H), 7.05 (s, 1H), 7.09 (s, 1H), 7.18 (s, 1H), 7.25 (s, 1H), 7.35 (s, 1H), 8.04 (broad s, 1H), 9.90 (s, 1H), 9.99 (s, 1H), 10.50 (s, 1H).

Bisdistamycin amine fumaramide 57. To a solution of 35 mg (0.027 mMol) **34** in 0.75 mL CH_2Cl_2 was added 0.5 mL TFA. The reaction was allowed to stir for 10 min and was precipitated with 3 mL anhydrous diethyl ether. The solid was filtered off, washed with anhydrous ether and dried under vacuum. It was then purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 3:25, v/v) to afford 24 mg (82%) **57**. ^{1}H NMR (DMSO-d_6) δ 1.59 (m, 2H), 2.11 (s, 3H), 2.30 (m, 4H), 3.18 (q, J=8 Hz, 2H), 3.23 (q, J=5 Hz, 2H), 3.80 (s, 3H), 3.85 (s, 3H), 3.88 (s, 3H), 6.84 (s, 1H), 6.98 (s, 1H), 7.05 (s, 1H), 7.10 (s, 1H), 7.20 (s, 1H), 7.27 (s, 1H), 7.34 (s, 1H), 8.04 (t, J=1H), 9.91 (s, 1H), 9.99 (s, 1H), 10.51 (s, 1H); m/e 1104 ($M+H^+$).

Bis-EDTA-Distamycin Fumaramide-Hexa Ethyl Ester 35. A solution of 92 mg (0.24 mMol) EDTA triethylester in 5 mL DMF was activated as

the imidazolide by the addition of 40 mg (0.25 mMol) *N,N'*-carbonyldiimidazole and stirring 2 h. After this time, the activated EDTA triethylester was added to a solution of 36.25 mg (0.0329 mMol) **57** in 2 mL DMF and the coupling reaction was allowed to proceed for 12 h. The reaction was then concentrated under vacuum and triturated three times with diethyl ether. The sample was flash chromatagraphed twice ((concentrated aqueous ammonia:(ethyl acetate:ethanol, 3:2, v/v), 3:97, v/v) then (concentrated aqueous ammonia:methanol, 1:99, v/v)) to produce 24 mg (40%) **35**. ^{1}H NMR (DMSO-d_6) δ 1.18 (t, J=7 Hz, 9H), 1.56 (m, 2H), 1.62 (m, 2H), 2.13 (s, 3H), 2.30 (m, 4H), 2.71 (m, 4H), 3.12 (q, J=6 Hz, 2H), 3.17 (m, 4H), 3.50 (s, 2H), 3.59 (s, 4H), 3.80 (s, 3H), 3.85 (s, 3H), 3.88 (s, 3H), 4.05 (q, J=7 Hz, 6H), 6.84 (s, 1H), 6.96 (s, 1H), 7.08 (s, 1H), 7.10 (s, 1H), 7.20 (s, 1H), 7.27 (s, 1H), 7.36 (s, 1H), 7.99 (t, J=7 Hz, 1H), 8.05 (t, J=7 Hz, 1H), 9.90 (s, 1H), 9.99 (s, 1H), 10.50 (s, 1H); m/e 1820 ($M+H^+$).

Bis-EDTA-Distamycin Fumaramide 31. A solution of 16.8 mg (0.0092 mMol) bis-EDTA-distamycin fumaramide-hexa ethyl ester **35** in 1 mL ethanol and 1 mL water was hydrolyzed by the addition of 4 mg (0.167 mMol) lithium hydroxide. The hydrolysis reaction was allowed to proceed for 12 h when the reaction was concentrated under vacuum. The compound was purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 3:97, v/v) to yield 16.2 mg (106%) **31**. IR (KBr) 3010, 2920, 2850, 1735, 1640, 1600, 1440, 1400, 1260, 1200, 1140, 1100 cm^{-1}; ^{1}H NMR (DMSO-d_6) δ 1.90 (m, 4H), 2.85 (s, 3H), 3.13 (m, 2H), 3.22 (m, 2H), 3.28 (m, 2H), 3.32 (m, 2H), 3.40 (m,

2H), 3.46 (m, 2H), 3.85 (s, 3H), 3.90 (s, 3H), 3.93 (s, 3H), 3.95 (s, 4H), 4.00 (s, 2H), 4.10 (s, 2H), 6.96 (s, 1H), 7.00 (s, 1H), 7.11 (s, 2H), 7.18 (s, 1H), 7.37 (s, 1H), 8.13 (broad s, 1H) 8.50 (m, 1H), 9.98 (s, 1H), 10.03 (s, 1H), 10.54 (s, 1H); UV (H_2O) λ_{max} 241 nm, 310 nm; m/e 1652 ($M+H^+$).

***t*-Boc β-alanine 39.** A solution of 6.23 g (70 mMol) β-alanine and 2.5 g (70 mMol) sodium hydroxide in 210 mL dioxane:water (2:1, v/v) was cooled to 0°C in an ice water bath. Solid *t*-boc anhydride (16.8 g, 77 mMol) was added and allowed to stir at 0°C for 0.5 h and then room temperature for 1 h. The solution was rotovapped at 35°C until distillation had nearly stopped and 300 mL ethyl acetate was added. The pH of the solution was adjusted to pH 3 with 2 M potassium bisulfate. The mixture was shaken and the organic phase was separated. The aqueous layer was extracted two more times with ethyl acetate (200 mL) and the organic layers were combined. After washing once with water, the combined organic layers were dried over anhydrous sodium sulfate, the solvent was removed by rotoevaporation and the resulting solid was dried under high vacuum to yield 11.33 g (87%) *t*-boc β-alanine **39.** ^{1}H NMR (DMSO-d_6) δ 1.37 (s, 9H), 2.34 (t, J=7 Hz, 2H), 3.12 (q, J=7 Hz, 2H), 6.79 (br s, 1H); m/e 190 ($M+H^+$).

***t*-Boc β-alaninyl-tetra-*N*-methylpyrrolecarboxamide dimethylpropylamine 40.** A solution of 138 mg (0.22 mMol) 4-nitro-tetra-*N*-methylpyrrole-2-carboxamide dimethylpropylamine in 5 mL DMF was hydrogenated at atmospheric pressure over 5% palladium on charcoal (40 mg). The *t*-boc β-alanine

imidazolide was prepared by dissolving 62.5 mg (0.33 mMol) *t*-boc β-alanine in 2 mL DMF, adding 60 mg (0.37 mMol) *N,N'*-carbonyldiimidazole and stirring 1 h. The hydrogenated 4-nitro-tetra-*N*-methylpyrrole-2-carboxamide dimethylpropylamine was filtered through celite to remove catalyst, and added to the *t*-boc β-alanine imidazolide and stirred 12 h. The solvent was removed under high vacuum and the residue was triturated twice with diethyl ether and was purified by flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 2:98, v/v) to afford 114 mg (68%) **40**. ^{1}H NMR (CD_3CN) δ 1.40 (s, 9H), 1.67 (m, 2H), 2.23 (s, 6H), 2.37 (t, J=7 Hz, 2H), 2.44 (t, J=7 Hz, 2H), 3.32 (m, 4H), 3.86 (s, 3H), 3.87 (s, 3H), 3.89 (s, 6H), 5.48 (broad s, 1H), 6.61 (s, 1H), 6.76 (s, 1H), 6.89 (s, 1H), 6.91 (s, 1H), 7.14 (s, 1H), 7.21 (s, 3H), 7.53 (broad s, 1H), 8.30 (s, 1H), 8.40 (s, 2H), 8.44 (s, 1H).

***t*-Boc β-alaninyl-tetra-N-methylpyrrolecarboxamidecarboxylic acid methyl ester 42.** A solution of 138 mg (0.25 mMol) 4-nitro-tetra-*N*-methylpyrrole-2-carboxylic acid methyl ester in 5 mL DMF was hydrogenated over 5% palladium on charcoal (40 mg) at atmospheric pressure. *t*-Boc β-alanine was activated as the imidazolide by dissolving 70 mg (0.37 mMol) *t*-boc β-alanine in 1 mL DMF and adding 67 mg (0.41 mMol) *N,N'*-carbonyldiimidazole and stirring for 1 h. The hydrogenation mixture was filtered through Celite to remove the catalyst, added to the activated acid and the solution was allowed to stir 12 h. The solvent was removed under high vacuum, the residue was triturated with diethyl ether three times and purified by flash chromatography on silica

gel (methanol:dichloromethane, 1:9, v/v) to give 66 mg (38%) **42**. ^{1}H NMR (DMSO-d_6) δ 1.38 (s, 9H), 2.40 (t, J=7 Hz, 2H), 3.19 (q, J=7 Hz, 2H), 3.74 (s, 3H), 3.84 (s, 6H), 3.85 (s, 3H), 3.86 (s, 3H), 6.81 (t, J=5 Hz, 1H), 6.88 (d, J=1.7 Hz, 1H), 6.91 (d, J=1.3 Hz, 1H), 7.05 (s, 1H), 7.08 (d, J=1.7 Hz, 1H), 7.16 (s, 1H), 7.24 (s, 2H), 7.47 (d, J=1.7 Hz, 1H), 9.84 (s, 1H), 9.88 (s, 1H), 9.90 (s, 1H), 9.93 (s, 1H).

***t*-Boc β-alaninyl-tetra-N-methylpyrrolecarboxamidecarboxylic acid 43.** A suspension of 54 mg (0.078 mMol) *t*-boc β-alaninyl-tetra-*N*-methyl-pyrrolecarboxamidecarboxylic acid methyl ester **42** in 5 mL water and 5 mL ethanol was added 11.4 mg (0.29 mMol) sodium hydroxide and the mixture was refluxed for 3 h. At this time an additional 7 mg (0.18 mMol) sodium hydroxide was added and the solution was refluxed for another hour. After one hour, the solution was clear and the ethanol was rotovapped off. The solution was cooled in an ice bath and 0.6 mL 1 N HCL (0.6 mMol) was added in 5 mL ice cold water. The resulting mixture was lyophilized and washed three times with one mL water. The solid was isolated after each washing by centrifugation. The 43 mg resulting solid was purified by flash chromatography on silica gel with (concentrated aqueous ammonia:(methanol:dichloromethane, 1:1, v/v), 1:33, v/v) to afford 19 mg (36%) **43**. ^{1}H NMR (DMSO-d_6) δ 1.38 (s, 9H), 2.40 (t, J=7 Hz, 2H), 3.20 (q, J=7 Hz, 2H), 3.84 (s, 3H), 3.85 (s, 3H), 3.86 (s, 3H), 6.82 (t, J=1H), 6.89 (d, J=2 Hz, 1H), 7.06 (m, 2H), 7.11 (d, J=2 Hz, 1H), 7.25 (s, 1H), 7.69 (d, J=2 Hz, 1H), 9.86 (s, 1H), 9.92 (s, 1H), 9.96 (s, 1H), 10.02 (s, 1H); m/e 677 ($M+H^+$).

β-alaninyl-tetra-N-methylpyrrolecarboxamide dimethylpropylamine 41. A solution of 94 mg (0.136 mMol) **40** in 6 mL CH_2Cl_2 and 3 mL TFA was stirred for 10 min and the product was precipitated by the addition of 40 mL diethyl ether. The precipitate was purified by flash chromatography on silica gel with (concentrated aqueous ammonia:methanol, 11:89, v/v) to give 73 mg (81%) **41**. ^{1}H NMR (DMSO-d_6) δ 1.61 (p, J=7 Hz, 2H), 2.14 (s, 6H), 2.25 (t, J=7 Hz, 2H), 2.42 (t, J=7 Hz, 2H), 2.89 (t, J=7 Hz, 2H), 3.19 (q, J=6 Hz, 2H), 3.80 (s, 3H), 3.84 (s, 3H), 3.85 (s, 3H), 3.86 (s, 3H), 6.83 (d, J=2 Hz, 1H), 6.89 (d, J=2 Hz, 1H), 7.06 (m, 2H), 7.18 (m, 2H), 7.24 (m, 2H), 8.06 (t, J=5 Hz, 1H), 9.89 (t, J=5 Hz, 1H), 9.89 (s, 1H), 9.90 (s, 1H), 9.94 (s, 1H), 9.99 (s, 1H).

***t*-Boc β-alanine-tetra-N-methylpyrrolecarboxamide-β-alaninyl-tetra-N-methylpyrrolecarboxamide-dimethly propylamine 44.** A solution of 13.7 mg (0.021 mMol) β-alaninyl-tetra-*N*-methylpyrrolecarboxamide dimethyl propyl amine **41**, 15 mg (0.022 mMol) *t*-boc β-alaninyl-tetra-*N*-methylpyrrolecarboxamide carboxylic acid **43**, 5 mg (0.024 mMol) dicyclohexylcarbodiimide, and 6 mg (0.044 mMol) *N*-hydroxybenzotriazole in 5 mL DMF were stirred for 24 h. The DMF was removed under high vacuum and the residue was triturated twice with diethyl ether. After flash chromatography on silica gel (concentrated aqueous ammonia:methanol, 1:39, v/v), the fractions were allowed to stand and product precipitated. Combining this product with that isolated from the supernatant gave 26 mg (94%) **44**. ^{1}H NMR (DMSO-d_6) δ 1.38 (m, 9H), 2.40 (t, J=7 Hz, 4H), 3.20 (t, J=6 Hz, 4H), 3.80 (s, 3H), 3.82 (s, 3H), 3.83 (s, 3H), 3.85

(s, 3H), 6.84 (t, J=5 Hz, 1H), 6.86 (s, 2H), 6.89 (s, 2H), 7.06 (s, 4H), 7.16 (s, 1H), 7.18 (s, 1H), 7.20 (s, 2H), 7.23 (s, 4H), 8.09 (t, J=5 Hz, 1H), 8.10 (t, J=5 Hz, 1H), 9.87 (s, 1H), 9.92 (s, 5H), 9.94 (s, 2H); m/e 1322 ($M+H^+$).

β-alanine-tetra-N-methylpyrrolecarboxamide-β-alaninyl-tetra-N-methyl-pyrrolecarboxamide-dimethly propylamine 45. A solution of 5.1 mg (0.0042 mMol) **44** in 1 mL TFA:CH_2Cl_2 (2:1, v/v) was stirred for 15 min. and precipitated with anhydrous diethyl ether. The solid was triturated with (concentrated aqueous ammonia:methanol, 1:10, v/v) and washed with water and dried under high vacuum to afford 4.4 mg (93%) **45**. ^1H NMR (DMSO-d_6) δ 1.60 (p, J=6 Hz, 2H), 2.14 (s, 6H), 2.23 (t, J=7 Hz, 2H), 2.32 (m, 4H), 2.81 (t, J=2H), 3.25 (m, 4H), 3.80 (s, 3H), 3.82 (s, 3H), 3.85 (m, 18H), 6.83 (s, 1H), 6.86 (s, 1H), 6.89 (s, 2H), 7.06 (s, 4H), 7.16 (s, 1H), 7.18 (s, 1H), 7.20 (s, 2H), 7.25 (s, 4H), 8.07 (t, J=3 Hz, 2H), 9.91 (m, 8H).

$(P4)_3E$ triethyl ester 46. To a solution of 4.4 mg (0.0036 mMol) **45** in 0.5 mL DMF and 0.25 mL CH_2Cl_2 was added 3.6 mg (0.0038 mMol) **24**, 0.91 mg (0.0044 mMol) dicyclohexylcarbodiimide and 1.0 mg (0.0074 mMol) *N*-hydroxybenzotriazole. The mixture was allowed to stir 12 h. At this time, it was concentrated under vacuum and triturated with anhydrous diethyl ether once. The chromatography column was equilibrated with (concentrated aqueous ammonia:(methanol:dichloromethane, 1:1, v/v), 1:99, v/v) and the solvent was changed to 1:1 methanol:CH_2Cl_2. The crude reaction mixture was loaded onto the column in 1:1 methanol:CH_2Cl_2 (v/v) and three column volumes of solvent

were run through the column. The solvent system was changed to (concentrated aqueous ammonia:(methanol:dichloromethane, 1:1, v/v), 1:99, v/v) and the product was collected to afford 3.6 mg (46%) **46**. ^{1}H NMR (DMSO-d_6) δ 1.22 (m, 9H), 1.75 (p, J=7 Hz, 2H), 1.89 (p, J=7 Hz, 2H), 2.26 (s, 6H), 2.34 (t, J=7 Hz, 2H), 2.40 (t, J=7 Hz, 2H), 2.59 (t, J=6 Hz, 4H), 2.76 (broad s, 4H), 3.51 (s, 2H), 3.66 (s, 4H), 3.86 (m, 24H), 4.12 (m, 6H), 6.70 (m, 2.5H),[128] 6.88 (m, 5.5H),[128] 7.17 (m, 9H);[128] m/e 2153 ($M+H^+$).

$(P4)_3E$ 36. To a solution of 1.8 mg (0.00084 mMol) of **46** in 0.6 mL ethanol was added 0.5 mg (0.021 mMol) lithium hydroxide in 0.4 mL water. The reaction was allowed to stir for 16 h at which time it was concentrated to dryness. The sample was then neutralized at 0°C with HCl and lyophilized to afford **36**.

During the preparation of the NMR sample, a hydrolysis mixture of $(P4)_3E$ was taken up in 1:1 methanol/water. This solution was cooled and precipitate formed. The sample was then neutralized with 1 μL of TFA and copious precipitate formed. This precipitate was isolated and used for the NMR sample. Subsequent studies on the solubility of $(P4)_3E$ found that it was easily soluble in basic water but only sparingly soluble in methanol, neutral water, water/pyridine, or water/pyridine/DMF. IR (KBr) 3010, 2920, 2850, 1730, 1645, 1635, 1600, 1580, 1490, 1450, 1385, 1250, 1200, 1140 cm^{-1}; ^{1}H NMR (DMSO-d_6+TFA) δ 1.74 (m, 2H), 1.85 (m, 2H), 2.29 (t, J=7 Hz, 2H), 2.58 (m, 4H), 2.80 (d, J=5 Hz, 6H), 3.08 (m, 2H), 3.17 (m, 4H), 3.29 (m, 4H), 3.46 (m, 4H), 3.75 (s, 4H), 3.85 (m, 38H), 4.07 (s, 2H), 6.87 (s, 2H), 6.90 (s, 2H), 6.96 (s, 1H), 7.07

(m, 6H), 7.19 (m, 6H), 7.20 (s, 6H), 8.08 (t, 1H), 8.16 (t, 1H), 8.40 (t, 1H), 9.82 (s, 1H), 9.91 (m, 11H); UV (H_2O) λ_{max} 240 nm, 312 nm.

Protected Spermidine-EDTA 48. The imidazolide of EDTA triethyl ester **16** was prepared by adding 0.8 g (4.9 mMol) carbonyl diimidazole to a solution of 1.84 g (4.9 mMol) EDTA triethyl ester **16** in 12 mL DMF and stirring 2 h. The EDTA triethyl ester imidazolide was added to a solution of 1 g (4.9 mMol) spermine in 20 mL DMF and stirred 12 h. At this time, 3.5 g (16 mMol) di-*tert*-butyl dicarbonate was added and the solution was heated to 37°C and stirred for 4 h. After the DMF was distilled off under vacuum, the residue was taken up in 200 mL diethyl ether and was extracted three times with 150 mL water, once with saturated sodium chloride and then dried over sodium sulfate. After the removing the solvent under reduced pressure, the resulting oil was dissolved (acetonitrile/dichloromethane, 2:3, v/v), loaded on a silica gel column, and washed with 975 mL of the solvent system. The solvent system was changed (ethanol/ethyl acetate, 1:20, v/v) and the protected compound was eluted from the column to give 1.3 g (31%) **48**. ^{1}H NMR (DMSO-d_6) δ 1.18 (m, 9H), 1.38 (m, 27H), 1.57 (m, 4H), 2.69 (q, 2H, J=6Hz), 2.73 (m, 2H), 2.88 (q, 2H, J=6Hz), 3.1 (m, 10H), 3.17 (s, 2H), 3.45 (s, 2H), 3.51 (s, 4H), 4.07 (m, 6H), 6.76 (broad s, 1H), 7.99 (t, 1H, J=5Hz). Analysis for $C_{41}H_{76}N_6O_6$. Calculated C, 57.19; H, 9.00; N, 9.76. Found: C, 57.19; H, 8.86; N, 9.95. Exact Mass (pos. ion, FAB) for $C_{41}H_{77}N_6O_{13}^+$: calculated 861.5548; found 861.5531.

Spermidine EDTA 47. A solution of 60 mg (0.07 mMol) **48** in 5 mL denatured ethanol was hydrolyzed by the addition of 12 mg (0.5 mMol) lithium

hydroxide in 5 mL water and heating to 70°C for 3 h. The reaction mixture was dried under vacuum and taken up in 5 mL dichloromethane. To this solution, 1.5 mL TFA was added and the solution was stirred 12 min. The product was precipitated with diethyl ether and collected on a glass frit. After washing with diethyl ether and drying, 62 mg (85%) **47** was obtained as the TFA salt. ^{1}H NMR (DMSO-d_6+TFA) δ 1.62 (broad s, 4H), 1.76 (t, 2H, J=7Hz), 1.88 (m, 2H), 2.91 (m, 10H), 3.25 (m, 4H), 3.32 (m, 2H), 3.78 (s, 4H), 3.87 (s, 2H), 4.00 (s, 2H), 7.90 (broad s, 3H), 8.51 (t, 1H, J=5Hz), 8.65 (broad s, 2H), 8.76 (broad s, 2H). Exact Mass (pos. ion, FAB) for $C_{20}H_{41}N_6O_7^+$: calculated 477.3036; found 477.3041.

Plasmid Preparation

Cell Growth:

Prepare 2 L VB minimal media (1 L each in two 2 L flasks)

per 1 L:	
2 g	Casamino acids
20 mL	50X V-B salts
50 mL	autoclaved thymine (1 mg/mL)
H_2O to give ~980 mL total volume	

Autoclave on wet cycle, allow to cool to r.t., add:

20 mL	Sterile (autoclaved) 20% (w/v) D-glucose
2 mL	Ampicillin (25 mg/mL, filtered sterile 0.45 μM)
1 mL	Thiamine·HCl (1 mg/mL, filtered sterile 0.45 μM)

Inoculate 1 L of VB minimal w/4 mL of an overnight culture of bacteria in L-broth and 50 μg/mL ampicillin (to keep selective pressure on, increases the yield of plasmid). Shake at 37°C at 200 rpm.

Amplify when O.D.$_{590nm}$ = 0.6-0.9 (best to amplify at higher O.D.) Amplify by adding 2 mL Chloromycetin sodium succinate (100 mg/mL) per 1 L of broth (final concentration = 200 μg/mL)

Chloromycetin sodium succinate = Chloroamphenical (Cm) and the 100 mg/mL solution is made by adding 10 mL of sterile H_2O (*via* syringe) to a 1 g vial of Cm and shaking until it has completely dissolved.

Amplify for 12-17 hours at 37°C with shaking.

Harvest Cells:

Chill 2 L flasks on ice bath with frequent agitation

Load chilled cells into 500 mL bottles, balance with double pan balance.

Spin at 5000 rpm, 4°C for 20′, discard supernatants and wipe bottles dry.

Lysis:

Resuspend cell pellet in 4-6 mL of:
25% Sucrose
50 mM Tris pH 8 (less volume is better)

Transfer resuspended cells to a 70Ti (polycarbonate) tube and add:
1.6 mL lysozyme (5 mg/mL prepared just before use)
–Let sit 5′ on ice with occasional stirring
5 mL 0.5 M EDTA pH 8
–Let sit 5′ on ice with occasional stirring

One tube at a time, add:

10%	Triton X–100
50 mM	Tris pH 8
60 mM	EDTA pH 8

Solution to almost fill tube – (air bubble good for agitation)

Shake vigorously (!!) until the lysates "clear"
looks like shampoo that's been shaken, usually 20-30 sec is O.K.
try to avoid clumping of lysates

Balance tubes to ±0.01 g, load into Ti70 rotor
Spin 30 K rpm, 4°C for 30′

Pour off supernatants into 50 mL polypropylene centrifuge tube
** avoid gluey junk near cell debris **

Add 350 λ of 2 mg/mL RNase A (DNase free as in Maniatis), incubate 20′ at 37°C

Add 1/3 volume (~6-7 mL) of 30% PEG 8000/1.5M NaCl to each tube, shake to mix and incubate 30′-8 h on ice (long time = better yield).

Centrifuge PEG/NaCl mixture: 20′ at 6 K rpm, 4°C, pitch supernatants

Resuspend pellet in 2–5 mL cold 10X TE using a glass rod
100 mM Tris pH 8
10 mM EDTA pH 8

CsCl spins:

Set up 1st CsCl isopycnic banding run:
(use 1x40 mL quick seal tube per 500 mL of original broth)
59.52 g CsCl (ultrapure)
+ Pellets in 10X TE
+ 3.2 mL ethidium bromide (10 mg/mL)
+ 1X TE buffer to give a final weight of : 124.00 g

This equals 80 mL of 48% (w/w) CsCl and has a $\rho \sim 1.55$

Stir to dissolve CsCl
Load into VTi50 quickseal tubes, balance to ±0.01 g and seal
Run in VTi50 rotor at 42 K rpm, 17°C for 20 h

Vent tube at top with an 18 GA needle and pull lower (plasmid) band (using UV illumination) by puncturing quick seal tube ~1 cm below band with an 18 GA needle and 10 mL syringe. Keep needle bevel up and protect DNA from excessive exposure to UV light.

Transfer bands into VTi65 tubes, fill tubes with CsCl filling solution ($\rho = 1.55$, 400 μg/mL ethidium bromide)
Balance tubes to ±0.01 g, seal and load into VTi65.2 Rotor

Spin at 55 K rpm, 17°C for 12 h

Extract plasmid (lower) band as before, but use a 3 mL syringe

Extract ethidium bromide with 1X TE saturated isoamyl alcohol
about 12-15 extractions of 1 volume each are needed
(DNA is in bottom layer)
check for EtBr removal with UV light

Dialyze plasmid *vs.* 1X Reaction Buffer pH 7.9, 40 mM Tris acetate, 5 mM sodium acetate at 4°C.

High Resolution DNA Studies

Preparation of Specifically Labeled DNA Fragments. Endlabeling was carried out similarly for each fragment. The following is a description of the

procedure for the labeling and isolation of the 167bp and 517bp fragments from pBR322. Superhelical pBR322 plasmid DNA was digested with the restriction endonuclease *Eco*R I, then labeled at the 3′ end with (α^{32}P)dATP using the Klenow fragment of DNA polymerase I. A second digest with restriction endonuclease *Rsa* I yielded the 3′ labeled 167bp and 517bp fragments. These restriction fragments were isolated by polyacrylamide gel electrophoresis. Cleavage of pBR322 with *Eco*R I and successive treatment with calf alkaline phosphatase, (γ^{32}P)ATP and T4 polynucleotide kinase, and *Rsa* I yielded the 5′ labeled fragments. Guanine specific chemical cleavage lanes were carried out as according to Maxam and Gilbert.[129] The specific reaction procedures and conditions are outlined in the following description.

167/517 3′ and 5′ labelling of pBR322

*Eco*R I Digest of 15 μg pBR322
100 λ pBR322 (150 μg/ml)
20 λ 10x *Eco*R I buffer
78 λ DD H_2O
2 λ *Eco*R I (20 μ/λ)

200 λ at 37°C, >3 h

MINIGEL (1% agarose, 80V, 40-60 min)
1 λ digest
1 λ Ficol loading buffer
2 λ DD H_2O

4 λ load in 1 lane

Standard
2 λ pBR322 (150 μg/ml)
10 λ DD H_2O
4 λ Ficol loading buffer

16 λ load 4 λ/lane

Phenol Extract (2x)

200 λ Phenol (saturated with 1x reaction buffer)
vort
spin 20sec
discard LOWER layer

Ether Extract (3x)

900 λ Ether
vort
spin 20sec
discard UPPER layer

EtOH ppt. (2X)

20 λ 20% NaOAc, vort
600 λ EtOH, invert 5x
spin 20′
pull supers

200 λ 0.3M NaOAc, vort
600 λ EtOH, invert 5x
spin 20′
pull supers

EtOH Wash
200 λ 70% EtOH (cold)
spin 3′
pull supers

vac 3-10 min

CAPing for 5′ labelling

15 μg *Eco*R I'd pBR322
80 λ DD H_2O
10 λ 10x CAP Buffer
10 λ CAP (1 μ/λ)

100 λ 37°C, 30 min

Phenol Extract (2x)

100 λ Phenol (saturated with 1x reaction buffer)
vort
spin 20sec
discard LOWER layer

Chloroform Extract (1x, optional)

100 λ 24:1 $CHCl_3$ *i*-Amyl Alcohol
vort
spin 20sec
discard LOWER layer

Ether Extract (3x)

900 λ Ether
vort
spin 20sec
discard UPPER layer

EtOH ppt. (2x)

10 λ 20% NaOAc, vort
300 λ EtOH, invert 5x
spin 20′
pull supers

100 λ 0.3M NaOAc, vort
300 λ EtOH, invert 5x
spin 20′
pull supers

70% EtOH wash
200 λ 70% EtOH (cold)
spin 3′
vac 3-10 min

End Labelling 3′ and 5′

3′ Labelling

15 μg *Eco*R I'd pBR322
21 λ DD H_2O
5 λ 10x *Hae* III buffer
5 λ 66 mM DTT (10 mg/ml)
5 λ dTTP (10 mM)
10 λ alpha-^{32}PdATP
4 λ 5 u/λ polymerase

50 λ 25°C, 20 min

1 λ dATP (10 mM)
1 λ dTTP (10 mM)
1 λ dGTP (10 mM)
1 λ dCTP (10 mM)

52 λ 25°C, 10 min

5′ Labelling

15 μg *Eco*R I'd & CAP'd pBR322
34 λ DD H_2O
5 λ 10x Kinase buff
5 λ 66 mM DTT (10 mg/ml)
ca 4 λ γ-^{32}P dATP
2 λ Kinase (NBL, 2 μ/λ)

50 λ 37°C, 45 min

EtOH ppt. (2x)

10 λ 20% NaOAc, vort
300 λ EtOH, invert 5x
spin 20′
pull supers

100 λ 0.3M NaOAc, vort
300 λ EtOH, invert 5x
spin 20′
pull supers

70% EtOH wash
200 λ 70% EtOH (cold)
spin 3′
vac 3-10 min

RSA Digest

15 μg labelled DNA
32 λ DD H_2O
5 λ 10x RSA buffer
5 λ 66 mM DTT (10 mg/ml)
7.5 λ RSA (10 μ/λ)

50 λ 37°C, 1-3 h

EtOH ppt. (2x)

10 λ 20% NaOAc, vort
300 λ EtOH, invert 5x
spin 20′
pull supers

100 λ 0.3 M NaOAc, vort
300 λ EtOH, invert 5x
spin 20′
pull supers

70% EtOH wash
200 λ 70% EtOH (cold)
spin 3′
vac 3-10 min

Prep Acrylamide Gel (5%, 1:30)

Plug (10%, 1:30)

2.5 g Acrylamide
85 mg Bisacrylamide
30 mg Ammonia Persulfate
5 ml 5x TBE

to 25 ml with DD H_2O

stir ca 1/2 h
millipore filter
40 λ TEMED

GEL (5%, 1:30)

4.0 g Acrylamide
136 mg Bisacrylamide
48 mg Ammonia Persulfate
16 ml 5x TBE

to 80 ml with DD H_2O

stir *ca.* 1/2 h
millipore filter

40 λ TEMED, let stand couple hours at least

load labelled nucleotides and electrophorese at 240 V

until BPB reaches the bottom of the gel.

Maxam-Gilbert G-Reaction

Start Buffer

500 λ 100 mM Na Cacodylate (pH 8.0)
100 λ 100 mM $MgCl_2$
10 λ 100 mM EDTA
390 λ DD H_2O

1000 λ

Stop Buffer

94 λ 4N NaOAc
18 λ Thioethanol
126 λ DD H_2O
2 λ tRNA (2 mg/ml)

---

250 λ

G-Reaction

200 λ Start Buffer
1.5 λ CT (1 mM bp)
5 λ ^{32}P labelled DNA

---

206 λ

cool to 0°C
1 λ Dimethyl Sulfate
20°C, 8 min

50 λ Stop Buffer, vort
750 λ EtOH, invert 5x
cool 15′
spin 7′
pull supers

250 λ 0.3 M NaOAc, vort
750 λ EtOH, invert 5x
cool 15′
spin 7′
pull supers

100 λ 70% EtOH
spin 3′
pull supers

100 λ 1 M Piperidine (99 λ Piperidine + 900 λ DD H_2O)
90°C, 30 min

freeze, lyophilize

10 λ DD H_2O
freeze, lyophilize

10 λ DD H_2O
freeze, lyophilize

resuspend in formamide loading buffer (at ca. 0.3mR/λ)

Affinity Cleavage Reactions. The concentration of each poly-*N*-methylpyrrolecarboxamide-EDTA compound was determined spectroscopically using the extinction coefficients measured for the acetyl compounds (table 5.1). The extinction coefficients for P8E and P9E were estimated from the series of acetyl compounds containing two to seven *N*-methylpyrrolecarboxamide units. The extinction coefficients for BEDF and $(P4)_3E$ were estimated as a linear combination of the extinction coefficients for the monomers. This assumption is supported by the extinction coefficients measured for the netropsin dimers.[100, 102] Each oligopeptide-EDTA·Fe(II) molecule was prepared fresh by mixing at high concentration (1-10 mM) with 1-2 equivalents of $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ and dilluting to the appropriate concentration.

Table 5.1: Extinction Coefficients for Affinity Cleaving Compounds

Compound	ϵ ($M^{-1}cm^{-1}$) ($\times 1000$)	Compound	est. ϵ ($M^{-1}cm^{-1}$) ($\times 1000$)
P2	ϵ_{235}=21.6, ϵ_{296}=22.8		
P3	ϵ_{237}=28.0, ϵ_{303}=33.0	DE	ϵ_{305}=33.0
P4	ϵ_{238}=30.0, ϵ_{308}=40.0	P4E	ϵ_{310}=40.0
P5	ϵ_{238}=35.0, ϵ_{310}=46.0	P5E	ϵ_{313}=46.0
P6	ϵ_{240}=38.0, ϵ_{314}=53.0	P6E	ϵ_{314}=53.0
P7	ϵ_{240}=45.0, ϵ_{315}=64.0	P7E	ϵ_{314}=64.0
		P8E	ϵ_{314}=72.0
		P9E	ϵ_{314}=80.0
		BEDF	ϵ_{310}=66.0
		$(P4)_3E$	ϵ_{312}=120.0

The cleavage reactions were run with >5000 cpm of ^{32}P endlabeled restriction fragment made up to a total concentration of 100 μM in base pairs (final

concentration) with sonicated calf thymus DNA. To 9 μL of DNA/buffer solution, 3 μL of oligopeptide-EDTA·Fe(II) complex was added. Each oligopeptide-EDTA·Fe(II) complex was then allowed to equilibrate with the buffered DNA solution. Equilibration was typically carried out for 1-2 hours at 25°C-65°C. Cleavage was initiated by the addition of 3 μL of a 25 mM dithiothreitol (DTT) solution to the reaction mixture to give a total reaction volume of 15 μl and a final concentration of 5 mM DTT. The reaction buffer was prepared to 40 mM Tris base and 5 mM sodium acetate and then adjusted to pH 7.9 with acetic acid. The reactions were typically run for 1-2 hours at 25°C-37°C and were terminated by freezing.

Footprinting Reactions. The footprinting reactions were run in a buffer of pH 7.4 containing 10 mM Tris and 50 mM NaCl and were made up to a final DNA concentration of 100 μM (in base pairs) with sonicated calf thymus DNA. The footprinting reagents (MPE·Fe(II), EDTA·Fe(II), SE·Fe(II)) were prepared by mixing at 1 mM with one equivalent of $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ and diluting immediately. To a 6 μL solution of DNA and buffer, 3 μL of the molecule to be footprinted at an appropriate concentration was added. This solution was allowed to equilibrate at 37°C for 1 h. Next, 3 μL of footprinting reagent (10 μM MPE·Fe(II) or 250 μM SE·Fe(II) or EDTA·Fe(II)) was added and allowed to equilibrate for 15 min at 37°C. The cleavage reaction was initiated by the addition of 3 μL of 15 mM DTT or 5 mM sodium ascorbate to give final concentrations of 3 mM DTT and 1 mM sodium ascorbate, respectively. Footprinting

reactions were allowed to continue for 15 min at 37°C and were terminated by freezing.

The MPE·Fe(II) footprinting reactions were also run at different temperatures. To compensate for the effect of the temperature change on the rate of cleavage, cleavage times were varied as follows: 4°C, 4 hours; 20°C, 1 hour; 37°C, 15 min; 55°C, 3 min.

$(P4)_3E$·In(III) Footprinting. The $(P4)_3E$·In(III) was prepared by mixing $(P4)_3E$ which had been dissolved in ~pH 11 DDW at a concentration of 1 mM with two equivalents of $InCl_3$ in ~pH 3 DDW. The $(P4)_3E$ solution was basified with NaOH and the $InCl_3$ solution was acidified with HCl. The $(P4)_3E$·In(III) was diluted to the appropriate concentration and equilibrated with the DNA for one hour at 65°C. The MPE·Fe(II) for these reactions was prepared by mixing MPE at 2 mM with 1.9 equivalents of $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ and diluting immediately.

After equilibration, the samples were cooled to room temperature and 3 mM DTT (final concentration) and 1:30 MPE·Fe(II) to base DNA pairs (3.3 μM final concentration) were added. The samples were vortexed and allowed to react at 37°C for 13 min. These conditions were optimized for a 205 bp DNA fragment and may overdigest a larger fragment. Cleavage was terminated by freezing.

High Resolution Gels. The lyophilized reactions were suspended in 4 μL of 1X TBE, 80% (vol/vol) formamide at pH 8.3 containing bromophenol blue and

xylene cylanol tracking dyes. Recrystalization of the formamide is important to remove any traces of formic acid which may cause depurination upon heat denaturation. Storing the TBE/formamide loading buffer for extended periods of time results in the formation of trace formic acid, however, the dye/formamide mixture alone appears to be stable when frozen. Gel buffer (TBE) consisted of 89 mM Tris base, 89 mM boric acid and 2 mM Na_2EDTA, pH 8.3. The gels were either 0.4 mm thick, 40 cm long or wedge-shaped 0.2-0.6 mm thick, 40 cm long and were 5-8% polyacrylamide (1:20 crosslinked), 50% urea. The gels were pre-electrophoresed for 0.5 hour and then the samples were heat denatured at 90°C for 90 s, cooled to 0°C, loaded onto the gel, and electrophoresed at 1300-2000 V. Gels were typically run until either the bromophenol blue or xylene cylanol tracking dye reached the bottom of the gel. The gels could be dried using a BIO-RAD Model 483 Slab Gel Dryer. They were then exposed to Kodak X-Omat photographic film at -70°C with a duPont Cronex Lighting Plus intensification screen for 8-12 hours. The gels were subsequently re-exposed at -70°C without an intensification screen (dried gels could be exposed at room temperature). The time for the second exposure was determined by assuming the intensification screen decreased the exposure time by 10×. The gel was scanned and the data were recorded as a plot of absorbance versus distance scanned, quantitated and converted into histogram form.

Double Strand Cleavage

DNA Preparation. 10 μgs of pBR322 was linearized with *Sty* I. This enzyme recognizes an asymmetrical site on pBR322 at 1369, $5'-\overset{C}{G}\overset{C}{G}\overset{A}{T}\overset{A}{T}\overset{G}{C}\overset{G}{C}-3'$. Both ends

of the linearized plasmid were labeled with hot CTP and one end was selectively labeled with hot ATP using the 3′ endlabeling techniques described earlier. The DNA was purified on a 0.7% low melt agarose gel containing 0.25 μg/mL ethidium bromide (EtBr) run at 15 mA (~25 V) at 4°C overnight. The DNA was visualized by UV fluorescence of the EtBr and the bands were excized. To each of the samples, 800λ of 1X TE buffer was added and the sample was heated for 5 minutes at 65°C with vortexing at 3 minutes. The DNA was isolated from the agarose by extraction with: 500λ phenol (centrifuging at least 10 minutes to separate the phases); 250λ phenol and 250λ $CHCl_3$/*i*-amyl alchohol (24:1 v/v); 500λ $CHCl_3$. After these extractions, the sample was transfered to a 2.2 mL ependorf tube and extracted 5 times with anhydrous butanol to remove EtBr and to concentrate the sample. It was then transfered to a 1.5 mL ependorf tube and extracted once with anhydrous butanol and lyophilized (extracting once with $CHCl_3$ before lyophilization helps remove excess butanol). The DNA was resuspended with 50 λ DDW and precipitated with 25 λ, 7.5 M ammonium acetate and 150λ ethanol. After precipitation, the sample was washed once with 70% ethanol, dried, resuspended in 1X reaction buffer and the yield was determined by scintilation counting.

Agarose Gels. The DNA cleaving reactions were carried out as described for the high resolution gels. Equilibration was typically carried out at 65°C for 1 hour. The reactions were initiated by the addition of a DTT solution and carried out at 37°C for 2 hours. The activity loaded per lane was ~5000 cpm for the

pBR322 labeled at one end and ~10,000 cpm for DNA labeled at both ends. The final reaction volumes were typically 15 λ and to this reaction mixture was added 2 λ 10X ficol loading buffer. 10X Ficol loading buffer was made up as follows:

0.25% BPB	40 mg
0.25% XC	40 mg
25% ficol	3.75 gm
	to 15 mL

The samples were loaded onto a vertical 4 mm, 1% agarose gel and electrophoresed at 120 V until the bromophenol blue tracking dye reached the bottom of the gel. The gels were dried on a gel drier at 60°C and were exposed to Kodak X-Omat photographic film without an intensification screen overnight.

Time course studies were run at 10X scale and aliquots were quenched by freezing in dry ice. Loading buffer was added to the sample, the sample was thawed and immediately loaded onto the gel. The equilibration studies were run by preparing a large scale reaction mixture and aliquoting it into each of the reaction tubes to ensure that each reaction mixture was identical.

The gels with non-radioactive DNA were run under the same reaction conditions with 0.25 μg of DNA per reaction and a final reaction volume of 10 λ. The DNA was visualized by staining with ethidium bromide (0.25 μg/mL) and the gel was photographed with Polaroid type 55 positive/negative film.

The molecular weight marker lane for the radiolabeled DNA was produced by four individual digestions of double labeled pBR322. The general digestion procedure is outlined in table 5.2.

Table 5.2

10 λ	λ phage (1 μg) in TE pH 8.0
10 λ	both endlabeled hot pBR322
7 λ	9 λ to 100 λ diluted mercaptoethanol
10 λ	10X RE buffer
70 λ	DDW
2 λ	Restriction Enzyme
109 λ	Total Volume

After digestion, the reactions were ethanol precipitated once, counted and combined to produce the molecular weight marker lane. The enzymes used in these digestions and the molecular weights of the radiolabeled fragments produced with each enzyme upon digestion of *Sty* I linearized pBR 322 are given in table 5.3.

Table 5.3

*Eco*R I	1371; 2992
*Bam*H I	994; 3369
Xmn I	662; 1769
Pvu I	1897; 2466

The autoradiograms were scanned with an LKB densitometer and quantitated with the accompanying data analysis software. The molecular weights of the DNA fragments from the cleavage reactions were determined by comparison of the mobility of the bands observed with the mobility of the molecular weight standard. To obtain greater accuracy in molecular weight determination, some cleavage reactions were spiked with the molecular weight standard. This allowed direct comparison of the mobilities of the reaction fragments with the molecular weight standard and avoided error introduced by lane to lane variations in electrophoretic mobility.

Because each cleavage site produces two fragments of DNA which sum to 4365, the cleavage of double labeled pBR322 should produce a symmetrical banding pattern. The bands resulting from cleavage at each site can then be paired by their integration because equal amounts of each fragment should be produced. For any given difference in the length of two DNA fragments, the relative difference in migration will be greater for smaller fragments. Therefore, the lower molecular weight bands were considered to be more accurate. The degree to which the bands in the molecular weight standard deviate from a straight line provides a limit for the accuracy of molecular weight assignment by this procedure. Therefore, binding site assignments were checked by determining the correlation coefficient of the line produced by plotting the observed band migration *vs.* the log of the assigned molecular weight and comparing that result to the correlation coefficient determined for the molecular weight marker lane.

λ-phage Gels The DNA cleavage reactions were run as described earlier for the high resolution gels with 0.25 μg of DNA per reaction and a final reaction volume of 10 λ (~100 μM in base pairs). Resuspending precipitated λ-phage DNA was very difficult. It required high temperatures and alot of vortexing. It may be easier to remove EDTA from the DNA by dialysis. After EDTA removal, the concentration of the DNA was determined by UV absorbance at 260 nm using the relationship 1A=45.45 μg/mL. Equilibration was typically carried out at 65°C for 2 hours. The reactions were initiated by the addition of a DTT solution to a final concentration of 5 mM and cleavage was carried

out at 37°C for 2 hours. At this time, 2 λ 10X ficol loading buffer was added to each reaction and they were loaded onto a 0.3% non-denaturing agarose gel containing 0.25 μg/mL ethidium bromide. The gel was run in the cold room (4°C) for 12-20 h with the buffer recirculated by a peristaltic pump until the XC marker dye was about 2-3 inches from the bottom of the gel. It is important to run these gels in the cold room because at room temperature they have the consistency of warm jello. Recirculating the buffer is necessary because the long electrophoresis times cause a pH gradient to be set up which becomes great enough to destroy the gel. After electrophoresis, the gels were photographed with Polaroid type 55 positive/negative film or Polaroid type 667 film.

References

[1] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J. D. Watson, *Molecular Biology of The Cell*, Garland, New York, 1983.

[2] J. Darnell, H. Lodish, D. Baltimore, *Molecular Cell Biology*, Scientific American, United States, 1986.

[3] E. F. Gale, E. Cundliffe, P. E. Reynolds, M. H. Richmond, M. J. Waring,*The Molecular Basis of Antibiotic Action*,John Wiley & Sons,London,1981.

[4] F. Arcamone, F. Bizioli, G. Canevazzi, A. Grein, German Pat. #1,027,667 (1958).

[5] A. Finlay, F. Hochstein, B. Sobin, F. Murphy, *J. Am. Chem. Soc.*, **73**, 342 (1951).

[6] F. E. Hahn in *Antibiotics III. Mechanisms of Action of Antimicrobial and Antitumor Agents* D. Gottlieb, P. D. Shaw, J. W. Corcoran, Eds., Springer, New York, 79-100 (1975).

[7] Ch. Zimmer, *Progress in Nucleic Acids Research and Molecular Biology*, **15**, 285 (1980).

[8] A. Krey, *Prog. in Mol. Subcell. Biol.*, **7**, 43 (1980).

[9] B. Puschendorf, E. Peterson, H. Wolf, H. Werchau, H. Grunicke, *Biochim. Biophys. Res. Comm.*, **43**, 617 (1971).

[10] F. Arcamone, F. Bizioli, G. Canevazzi, A. Grein, *Nature*, **203**, 1064 (1964).

[11] M. Julia, N. Preau-Joseph, *Compt. Rend.*, **257**, 1115 (1963).

[12] S. Penco, S. Redaelli, F. Arcamone, *Gazzetta Chimica Italiana*, **97**, 1110 (1967).

[13] M. Julia, N. Preau-Joseph, *C. R. Hebd. Seances Acad. Sci.*, **247**, 1115 (1963).

[14] H. M. Berman, S. Neidle, C. Zimmer, H. Thrum, *Biochim. et Biophys. Acta*, **561**, 124-131 (1979).

[15] G. V. Gurskaya, S. L. Grokhovsky, A. L. Zhuze, B. P. Gottikh, *Biochim. et Biophys. Acta*, **563**, 336-342 (1979).

[16] Ch. Zimmer, U. Waehnert, *Prog. Biophys. Molec. Biol.*, **47**, 31-112 (1986).

[17] G. Luck, Ch. Zimmer, *Studia Biophysica*, **40**, 9 (1973).

[18] C. Minchenkova, C. Zimmer, *Biopolymers*, **19**, 823 (1980).

[19] Ch. Zimmer, N. Kakiuchi, W. Guschlbauer, *Nucl. Acids Res.*, **10**, 1721-1731 (1982).

[20] C. Marck, N. Kakiuchi, W. Guschlbauer, *Nucl. Acids Res.*, **10**, 6147 (1982).

[21] R. Wartell, J. Larson, R. Wells, *J. Biol. Chem.*, **219**, 6719 (1974).

[22] Ch. Zimmer, K. Reinert, G. Luck, U. Waehnert, G. Lober, H. J. Thrum, *J. Mol. Biol.*, **58**, 329 (1971).

[23] K. Reinert, *J. Mol. Biol.*, **72**, 593 (1972).

[24] K. Reinert, *Nucl. Acids Res.*, **7**, 1375 (1979).

[25] K. Reinert, *Nucl. Acids Res.*, **8**, 5519 (1980).

[26] G. Luck, Ch. Zimmer, K. Reinert, F. Arcomone, *Nucl. Acids Res.*, **4**, 2655 (1977).

[27] K. Reinert, *Biophys. Chem.*, **13**, 1 (1981).

[28] K. Reinert, *Stud. Biophys.*, **67**, 49 (1978).

[29] G. Luck, H. Triebel, M. Waring, Ch. Zimmer, *Nucl. Acids Res.*, **1**, 503 (1974).

[30] A. Malcolm, G. Snovnou, *Cold Spring Harbor Symposium on Quantitative Biology*, **47**, 323 (1982).

[31] J. Sutherland, J. Duval, K. Griffin, *Biochemistry*, **17**, 5088 (1980).

[32] Ch. Zimmer, C. Marck, C. Schneider, D. Thiele, G. Luck, W. Guschlbauer, *Biochim. et Biophys. Acta*, **607**, 232 (1980).

[33] N. Dattagupta, M. Hogan, D. Crothers, *Biochemistry*, **19**, 5998 (1980).

[34] M. Hogan, N. Dattagupta, D. Crothers, *Nature*, **278**, 521 (1979).

[35] H. Wu, D. H. Crothers, *Nature*, **308**, 509-513 (1984).

[36] M. L. Kopka, C. Yoon, D. Goodsell, P. Pjura, R. E. Dickerson, *J. Mol. Biol.*, **183**, 553-563 (1985).

[37] M. L. Kopka, C. Yoon, D. Goodsell, P. Pjura, R. E. Dickerson, *Proc. Natl. Acad.*

Sci., U.S.A., **82**, 1376-1380 (1985).

[38] M. L. Kopka, P. Pjura, C. Yoon, D. Goodsell, R. E. Dickerson in *Structure & Motion: Membranes, Nucleic Acids & Proteins*, E. Clementi, G. Corongiu, M. H. Sarma and R. H. Sarma, Eds., Adenine Press, 1985.

[39] M. H. Sarma, G. Gupta, R. Sarma, *J. Biomol. Struc. Dyn.*, **3**, 433-436 (1985).

[40] B. R. Reid, lecture at the Industrial Associates Conference, California Institute of Technology, Pasadena, California, April 1, 1987.

[41] M. W. Van Dyke, R. P. Hertzberg, P. B. Dervan, *Proc. Natl. Acad. Sci., U.S.A.*, **79**, 5470 (1982).

[42] M. W. Van Dyke, P. B. Dervan, *Cold Spring Harbor Symposium on Quantitative Biology*, **47**, 347 (1983).

[43] M. W. Van Dyke, P. B. Dervan, *Nucl. Acids Res.*, **11**, 5555 (1983).

[44] M. J. Lane, J. C. Dabrowiak, J. N. Vournakis, *Proc. Natl. Acad. Sci., U.S.A.*, **80**, 3260-3264 (1983).

[45] K. R. Fox, M. J. Waring, *Nucl. Acids Res.*, **24**, 9271-9285 (1984).

[46] K. H. Harshman, P. B. Dervan, *Nucl. Acids Res.*, **13**, 4825-4835 (1985).

[47] J. W. Lown, K. Krowicki, U. G. Bhat, A. Skorogaty, B. Ward, J. C. Dabrowiak, *Biochemistry*, **25**, 7408-7416 (1986).

[48] D. J. Patel, L. L. Canuel, *Proc. Natl. Acad. Sci., U.S.A.*, **74**, 5207-5211 (1977).

[49] A. Pardi, K. M. Morden, D. J. Patel, I. Tinoco, Jr., *Biochemistry*, **22**, 1107-1113 (1983).

[50] D. J. Patel, *Proc. Natl. Acad. Sci., U.S.A.*, **79**, 6424-6428 (1982).

[51] D. J. Patel, *Eur. J. Biochem.*, **99**, 396-378 (1979).

[52] D. J. Patel, L. Shapiro, *Biochemie*, **67**, 887-915 (1985).

[53] D. J. Patel, L. Shapiro, *Biopolymers*, **25**, 707-727 (1986).

[54] D. J. Patel, L. Shapiro, *J. Biol. Chem.*, **261**, 1230-1240 (1986).

[55] R. E. Klevit, D. E. Wemmer, B. R. Reid, *Biochemistry*, **25**, 3296-3303 (1986).

[56] R. Wartell, J. Larson, R. Wells, *J. Biol. Chem.*, **219**, 6719 (1974).

[57] J. McGhee, *Biopolymers*, **15**, 1345 (1976).

[58] D. M. Crothers, *Biopolymers*, **10**, 2147 (1971).

[59] G. Luck, H. Triebel, M. Waring, Ch. Zimmer, *Nucl. Acids Res.*, **1**, 503 (1974).

[60] G. Mazza, A. Galizzi, A. Minghetti, A. Siccardi, *Antimicrob. Ag. Chemother.*, **3**, 384 (1973).

[61] A. S. Krylov, S. L. Grokhovsky, A. S. Zasedeatelev, A. L. Zhuze, G. V. Gursky, B. P. Gottikh, *Nucl. Acids Res.*, **6**, 289-304 (1979).

[62] G. V. Gursky, A. S. Zasedatelev, A. L. Zhuze, A. A. Khorlin, S. L. Grokhovsky, S. A. Streltsov, A. N. Surovaya, S. M. Nikitin, A. S. Krylov, V. O. Retchinsky, M. V. Mikhailov, R. S. Beabealashvili, B. P. Gottikh, *Cold Spring Harbor Symposium on Quantitative Biology*, **74**, 367-378 (1982).

[63] M. Waring, L. Wakelin, J. Lee, *Biochim. et Biophys. Acta*, **407**, 200 (1975).

[64] L. A. Marky, J. Curry, K. J. Breslauer, *Prog. Clin. Biol. Res.*, **172**, 155-173 (1985).

[65] L. A. Marky, K. S. Blumenfeld, K. J. Breslauer, *Nucl. Acids Res.*, **11**, 2857-2870 (1983).

[66] R. P. Hertzberg, P. B. Dervan, *J. Am. Chem. Soc.*, **104**, 313 (1982).

[67] R. P. Hertzberg, P. B. Dervan, *Biochemistry*, **23**, 3934 (1984).

[68] L. M. Chan, J. A. McCarter, *Biochim. et Biophys. Acta*, **204**, 252 (1970).

[69] P. G. Schultz, J. S. Taylor, P. B. Dervan, *J. Am. Chem. Soc.*, **104**, 6861 (1982).

[70] J. S. Taylor, P. G. Schultz, P. B. Dervan, *Tetrahedron*, **40**, 457 (1984).

[71] P. G. Schultz, Ph.D. Thesis, California Institute of Technology, Pasadena, California, 1984.

[72] W. D. Henner, S. M. Grunberg, W. A. Haseltine, *J. Biol. Chem.*, **257**, 11750-11754 (1982).

[73] J. F. Ward, M. M. Urist, *Int. J. Radiat. Biol.*, **12**, 209 (1967).

[74] R. Breslow, L. N. Lukens, *J. Biol. Chem.*, **235**, 292-296 (1960).

[75] J. E. Repine, O. W. Pfenninger, D. W. Talmage, E. M. Berger, D. J. Pettijohn, *Proc. Natl. Acad. Sci., U.S.A.*, **78**, 1001-1003 (1981).

[76] R. S. Youngquist, P. B. Dervan, *Nucl. Acids Res.*, to be submitted.

[77] J. M. Kean, Susan A. White, David E. Draper, *Biochemistry*, **24**, 5062-5070 (1985).

[78] F. Hutchinson, *Progress in Nucleic Acid Research and Molecular Biology*, **32**, 115-154 (1985).

[79] I. Saito, T. Morii, T. Matsuura, *Nucl. Acids Res. Symposium Series*, **12**, 95-98 (1983).

[80] S. Murai, N. Sonada, S. Tsutsumi, *Bull. Chem. Soc. Japan*, **36**, 527 (1963).

[81] R. V. Digman, D. F. Anderson, *J. Org. Chem.*, **28**, 239 (1963).

[82] F. Arcamone, V. Nicolella, S. Penco, S. Redaelli, *Gazz. Chim. Ital.*, **99**, 632 (1969).

[83] F. Zunino, A. DiMarco, *Biochemical Pharmacology*, **21**, 867 (1972).

[84] P. Chandra, A. DiMarco, F. Zunino, A. Casazza, D. Gericke, F. Giuliani, C. Soranzo, R. Thorbeck, A. Gotz, F. Arcamone, M. Ghione, *Naturwissenshaften*, **59**, 448 (1972).

[85] Ch. Zimmer, G. Luck, *FEBS Lett.*, **10**, 339 (1970).

[86] C. Zimmer, G. Luck, E. Birch-Hirschfeld, R. Weiss, F. Arcamone, W. Guschlbauer, *Biochim. et Biophys. Acta*, **741**, 15-22 (1983).

[87] P. G. Schultz, P. G. Dervan, *Proc. Natl. Acad. Sci., U.S.A.*, **80**, 6834 (1983).

[88] P. G. Schultz, P. G. Dervan, *J. Biomol. Struc. Dyn.*, **1**, 1133 (1984).

[89] D. Goodsell, R. E. Dickerson, *J. Med. Chem.*, **29**, 727 (1986).

[90] M. Bialer, G. Yagen, R. Mechoulam, *Tetrahedron*, **34**, 2389 (1979).

[91] R. Hay, K. Nolan, *J. Chem. Soc., Dalton Trans.*, 1348 (1975).

[92] Plasmid SV-CAT DNA was graciously provided by Dr. Nevis Fregien.

[93] K. R. Fox, M. J. Waring, *Nucl. Acids Res.*, **24**, 9271-9285 (1984).

[94] T. D. Tullius, B. A. Dombroski, *Proc. Natl. Acad. Sci., U.S.A.*, **83**, 5469-5473 (1986).

[95] S. B. Zimmerman, *Ann. Rev. Biochem.*, **51**, 395-427 (1982).

[96] D. Freifelder, B. Trumbo, *Biopolymers*, **7**, 681-693 (1969).

[97] The value is 4367 not 4363 because pBR322 has been linearized with *Sty* I and filled in with polymerase. This adds four base pairs to the total length of the DNA.

[98] One site, 3240-3257, contains two runs of $\geq$8 base pairs of A·T DNA which are not resolved on the gels and; therefore, will be considered one site for the purposes of this discussion.

[99] Yu. D. Nechipurenko, A. S. Krylov, A. S. Zasedatelev, G. V. Gurskii, *Mol. Biol.*, **18**, 263-272 (1984).

[100] A. A. Khorlin, A. S. Krylov, S. L. Grokhovsky, A. L. Zhuze, A. S. Zasedatelev, G. V. Gursky, *FEBS Letters*, **118**, 311-314 (1980).

[101] P. G. Schultz, P. G. Dervan, *J. Am. Chem. Soc.*, **105**, 7748 (1983).

[102] A. V. Skamrov, I. N. Rybalkin, R. Sh. Biblashvili, B. P. Gottikh, S. L. Grokhovskii, G. V. Gurskii, A. L. Zhuze, A. S. Zasedatelev, Yu. D. Nechipurenko, A. A. Khorlin, *Mol. Biol.*, **19**, 153-167 (1985).

[103] R. Scott Youngquist, Peter B. Dervan, *J. Am. Chem. Soc.*, **107**, 5528-5529 (1985).

[104] J. H. Griffin, P. B. Dervan unpublished observations.

[105] H. Moser, P. B. Dervan, *Science*, submitted.

[106] T. D. Tullius, B. A. Dombroski, *Proc. Natl. Acad. Sci., U.S.A.*, **83**, 5469-5473 (1986).

[107] H. R. Drew, A. A. Travers, *Cell*, **37**, 491-502 (1984).

[108] C. W. Tabor, H. Tabor, *Ann. Rev. Biochem.*, **53**, 749-790 (1984).

[109] S. Z. Hirschman, M. Leng, G. Felsenfeld, *Biopolymers*, **5**, 227-233 (1967).

[110] C. Bull, G. J. McClune, J. A. Fee, *J. Am. Chem. Soc.*, **105**, 5290-5300 (1983).

[111] W. D. Wilson, Y.-H. Wang, S. Kusuma, S. Chandrasekaran, N. C. Yang, D. W. Boykin, *J. Am. Chem. Soc.*, **107**, 4989-4995 (1985).

[112] R. S. Youngquist, P. B. Dervan, *Proc. Natl. Acad. Sci., U.S.A.*, **82**, 2565-2569 (1985).

[113] J. Brosius, R. L. Cate, A. P. Perlmutter, *J. Biol. Chem.*, **257**, 9205-9210 (1982).

[114] J. K. Barton, A. L. Raphael, *Proc. Natl. Acad. Sci., U.S.A.*, **82**, 6460-6464 (1985).

[115] D. M. J. Lilley, *Proc. Natl. Acad. Sci., U.S.A.*, **77**, 6468-6472 (1980).

[116] D. M. J. Lilley, E. Palecek, *EMBO*, **3**, 1187-1192 (1984).

[117] N. Panayotatos, R. D. Wells, *Nature*, **289**, 466-470 (1981).

[118] L. G. Sheflin, D. Kowalski, *Nucl. Acids Res.*, **13**, 6137-6154 (1985).

[119] J. C. Furlong, D. M. J. Lilley, *Nucl. Acids Res.*, **14**, 3995-4007 (1986).

[120] D. Mendel, P. B. Dervan, *Proc. Natl. Acad. Sci., U.S.A.*, **84**, 910-914 (1987).

[121] B. H. Johnston, A. Rich, *Cell*, **42**, 713-724 (1985).

[122] E. Selsing, R. D. Wells, C. J. Alden, S. Arnott, *J. Biol. Chem.*, **254**, 5417 (1978).

[123] H.-S. Koo, H.-M. Wu, D. M. Crothers, *Nature*, **320**, 501-506 (1986).

[124] P. J. Hagerman, *Biochemistry*, **24**, 7033-7037 (1985).

[125] J. G. Sutcliffe, *Cold Spring Harbor Symposium on Quantitative Biology*, **43**, 77-90 (1979).

[126] K. Peden, *Gene*, **22**, 277-280 (1983).

[127] T. Tanka, B. Weisblum, *J. Bacteriol.*, **121**, 354 (1979).

[128] These pyrrole peaks may have integrated low due to incomplete relaxation or because of H-D exchange with solvent. The integration for the pyrrole protons in d_6-DMSO was correct. Additionally, in DMSO the amide peaks were visible; 8.03 (t, 1H), 8.07 (broad s, 3H), 9.15 (m, 12H).

[129] A. M. Maxam, W. Gilbert, *Methods in Enzymology*, **65**, 499-560 (1980).